

STIC-Biotech/ChemLib

From: Mertz, Prema
Sent: Tuesday, June 04, 2002 12:09 PM
To: STIC-Biotech/ChemLib
Subject: 09/828,217

Please search SEQ ID NO:1 with protein databases.

Thanks
Prema Mertz, Ph.D.
Art Unit 1646
Mailbox 10C-01
Crystal Mall 1, Room 10E-01
United States Patent & Trademark Office
(703) 308-4229

CRFE

Searcher: AB Wesen - July
Phone: 308-4501/4508
Location: Biotech Lab.
Date Picked Up: 6/4/02
Date Completed: 6/5/02
Searcher Prep/Review: _____
Clerical: 3 m
Online time: 2 m

TYPE OF SEARCH:

NA Sequences: _____
AA Sequences: 1
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST(where applic.)

STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: AB5502
WWW/Internet: _____
Other (specify): _____

THIS PAGE BLANK (USPTO)

Db 575 GAADLQNRHARLEAESSGLTRRLQESAAYALRRKEDLERRTPAAS----AAQDVS 627

RESULT 2

AC 09K6X4 PRELIMINARY; PRT; 461 AA.

DT 01-OCT-2000 (TREMblrel. 15, Created)

DE 01-DEC-2001 (TREMblrel. 19, Last sequence update)

DE CELL WALL-BINDING PROTEIN.

GN B3600.

OC Bacillus halodurans.

OC Bacteria; Firmicutes; Bacillus/Clostridium group;

OC Bacillus/Staphylococcus group; Bacillus.

OX NCBI_TaxID=86665;

RN [1]

SEQUENCE FROM N.A.

STRAIN=C-125 / JCM 9153;

MEDLINE=20512582; PubMed=11058132;

RA Takami H., Nakasone K., Takaki Y.,

RA Fujii F., Hirama C., Nakamura Y.,

RA Horikoshi K.;

RT "Complete genome sequence of the alkaliphilic bacterium Bacillus

RT halodurans and genomic sequence comparison with Bacillus subtilis."

RL Nucleic Acids Res. 28:4317-4331(2000).

DR EMBL; AP001519; BAB07319.1;

DR MEROPS; M37.UPM;

DR InterPro: IPR002886; Peptidase_M37.

DR Pfam: PF01551; Peptidase_M37; 1.

KW Complete proteome.

SQ SEQUENCE 461 AA; 50372 MW; 2918480CD67AF3F CRC64;

Query Match 12.1%; Score 107.5; DB 16; Length 461;

Best Local Similarity 26.1%; Pred. No. 0.44; 55; Indels 5; Gaps 2;

Matches 31; Conservative 28; Mismatches 55; Indels 5; Gaps 2;

QY 33 IIVLVPLIIITIKANSEACRDLRAVMCRVTHLLOELTEAKGFQDVEAQAATCN 92

DB 5 ISVVAAGLTFILFSQSSIEDA-KANSSLQNGISDVQKREKQKQKTELEL----- 59

QY 93 HTYMALMASIDAEKAGQKQKVELEGITTLNKLQDASAEVERLRENOVLSVRIADK 151

DB 60 KEVEKELGDTAEIERLDKEVEETSGRIQEKREIEVEQAIEELKQIEILERRIAER 118

RESULT 3

Q9KA43

PRELIMINARY;

PRT; 143 AA.

AC 09KA43;

DT 01-OCT-2000 (TREMblrel. 15, Created)

DE 01-OCT-2000 (TREMblrel. 15, Last sequence update)

DE FLAGELLAR PROTEIN REQUIRED FOR FLAGELLAR FORMATION.

GN FLIL OR BH2447.

OS Bacillus halodurans.

OC Bacteria; Firmicutes; Bacillus/Clostridium group;

OC Bacillus/Staphylococcus group; Bacillus.

OX NCBI_TaxID=86665;

RN [1]

SEQUENCE FROM N.A.

STRAIN=C-125 / JCM 9153;

MEDLINE=20512582; PubMed=11058132;

RA Takami H., Nakasone K., Takaki Y.,

RA Fujii F., Hirama C., Nakamura Y.,

RA Horikoshi K.;

RT "Complete genome sequence of the alkaliphilic bacterium Bacillus

RT halodurans and genomic sequence comparison with Bacillus subtilis."

RL Nucleic Acids Res. 28:4317-4331(2000).

DR EMBL; AP001515; BAB06166.1;

DR Flagella; Complete proteome.

SQ SEQUENCE 143 AA; 16386 MW; 613446D64B1C5402 CRC64;

Query Match 12.0%; Score 107; DB 16; Length 143;

Best Local Similarity 22.1%; Pred. No. 0.13; 49; Indels 26; Gaps 4;

Matches 31; Conservative 34; Mismatches 49; Indels 26; Gaps 4;

QY 23 LIGGIVLVLLIVLVPLIIITIKANSEACRDLRAVMCRVTHLLOEL 74

DB 6 LVNMLILVLLIVLVPLIIITIKANSEACRDLRAVMCRVTHLLOEL 65

QY 75 TEA-----QKGFQDVEAQAATCNHYVMAASLDAEKAQKQKVELEGITTLNKH 126

DB 66 VRAEFLIHVDNRNALQEVKRDQFNQNIIRSLAGMASOLSGADGIEKLEAQ----- 118

QY 127 LODASAEVERLRENOVLSV 146

DB 119 LOD-----DINALMQEGSVYKI 135

RESULT 4

Q9Y6W2

PRELIMINARY; PRT; 756 AA.

AC 09Y6W2;

DT 01-NOV-1999 (TREMblrel. 12, Created)

DE 01-NOV-1999 (TREMblrel. 12, Last sequence update)

DE HCR PROTEIN.

GN HCR.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

SEQUENCE FROM N.A.

MEDLINE=20014706; PubMed=10545595;

RA Oka A., Tamaya G., Tomizawa M., Ota M., Katsuyama Y., Makino S.,

RA Shilina T., Yoshitome M., Lizuka M., Sasao Y., Iwashita K.,

RA Kawakubo Y., Sugai J., Ozawa A., Ohkido M., Kimura M., Bahram S.,

RA Inoko H.;

RT "Association analysis using refined microsatellite markers localizes a

RT susceptibility locus for psoriasis vulgaris within a 11kb segment

RL Hum. Mol. Genet. 8:2165-2170(1999).

DR EMBL; AB029331; BAA81890.1;

SQ SEQUENCE 756 AA; 86118 MW; 88FD5F858EF07601 CRC64;

Query Match 11.8%; Score 104.5; DB 4; Length 756;

Best Local Similarity 24.7%; Pred. No. 1.4; 48; Indels 37; Gaps 7;

Matches 39; Conservative 34; Mismatches 48; Indels 37; Gaps 7;

QY 54 RDLRLAVME-----CRNVTHL---QOELTEAKGFQDVEAQAATCN-----NHTVMA 97

DB 258 RDLRLAVME-----CRNVTHL---QOELTEAKGFQDVEAQAATCN-----NHTVMA 97

QY 98 LMASLDAEKAQKQKVELEGITTLNKH-----LODASAEVERLRENOV 143

DB 318 LMASLDAEKAQKQKVELEGITTLNKH-----LODASAEVERLRENOV 143

QY 144 LSVRIA-----DKKYRPSQDSSAAPOLLIVLGISA 177

DB 378 LSVRIA-----DKKYRPSQDSSAAPOLLIVLGISA 177

RESULT 5

Q9NRK8

PRELIMINARY; PRT; 756 AA.

AC 09NRK8;

DT 01-OCT-2000 (TREMblrel. 15, Created)

DE 01-DEC-2001 (TREMblrel. 19, Last sequence update)

DE A-HELICAL PROTEIN.

GN HCR.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-20347693; PubMed-10888604;
 RA Asunalahti K., Laitinen T., Itkonen-Vatjus R., Lokki M.-L.,
 RA Suomeala S., Stellan E., Saarialho-Kere U., Kere J.;
 RT "A candidate gene for psoriasis near HLA-C, HCR (P98), is highly
 RT polymorphic with a disease-associated susceptibility allele."
 RT Hum. Mol. Genet. 9:1533-1542(2000).
 DR EMBL: AF216493; AAF74221.1;
 SQ SEQUENCE 756 AA; 86058 MW; 8EAD03358B62DEB4 CRC64;

Query Match 11.8%; Score 104.5; DB 4; Length 756;
 Best Local Similarity 24.7%; Pred. No. 1.4;
 Matches 39; Conservative 34; Mismatches 48; Indels 37; Gaps 7;
 QY 54 RGLRAVME-----CRNTHLL---QOELTEAGKGFQDYEAQ-AAAC-----NHTVMA 97
 DB 258 RDSLHATAELQVRSVSLTHIALQEEELTRKVPSPDSLEPFRKQSLNMRREKVFA 317
 QY 98 LMASLDAEKAQKQKVEELEGETTTLNKH-----LQDASAEVERLRRENOV 143
 DB 318 LMVQLAKQLEHSDSVKQLKGVAISQEVYTSQSQEQAILQSLQDKAAEVEVERMGAKG 377
 QY 144 LSVRIA---DKKYPSSQDSSSAAPOLLIVLLGSA 177
 DB 378 LQELSLRAQEARRW---QQQTASAEQRLVYNAVSS 412

RESULT 6
 QY6W1 PRELIMINARY: PRT; 756 AA.
 AC QY6W1:
 DT 01-NOV-1999 (TREMBlrel. 12, Created)
 DT 01-NOV-1999 (TREMBlrel. 12, last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, last annotation update)
 DE HCR PROTEIN (TRICHOVALIN HOMOLOGUE).
 GN HCR.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-BLOOD;
 RA Oka A., Tamiya G., Makino S., Tomizawa M., Yamagata T., Shilina T.,
 RA Watanabe K., Yamazaki M., Tashiro H., Kimura M., Inoko H.;
 RT "HCR-a-helix coiled-coil rod homologue."
 RT Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Hirakawa M., Yamaguchi H., Imai K., Shimada J.;
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Shilina S., Tamiya G., Oka A., Inoko H.;
 RT "Homo sapiens 2,229,817bp genomic DNA of 6p21.3 HLA class I region."
 RT Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AB029343; BAA82158.1;
 DR EMBL: AF000509; BAB6313.1;
 SQ SEQUENCE 756 AA; 85951 MW; 3D109AF0EEBA9F5 CRC64;

Query Match 11.8%; Score 104.5; DB 4; Length 756;
 Best Local Similarity 25.2%; Pred. No. 1.4;
 Matches 40; Conservative 32; Mismatches 48; Indels 39; Gaps 7;
 QY 54 RGLRAVME-----CRNTHLL---QOELTEAGKGFQDYEAQ-AAAC-----NHTVMA 97

DB 258 RDSLHATAELQVRSVSLTHIALQEEELTRKVPSPDSLEPFRKQSLNMRREKVFA 317
 QY 98 LMASLDAEKAQKQKVEELEGETTTLNKH-----LQDASAEVERLRRENOV 143
 DB 318 LMVQLAKQLEHSDSVKQLKGVAISQEVYTSQSQEQAILQSLQDKAAEVEVERMGAKG 377
 QY 144 LSVRIA---DKKYPSSQDSSSAAPOLLIVLLGSA 177
 DB 378 LQELSLRAQEARRW---QQQTASAEQRLVYNAVSS 412

RESULT 7
 QY6XK3 PRELIMINARY: PRT; 782 AA.
 AC QY6XK3:
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, last sequence update)
 DT 01-OCT-2000 (TREMBlrel. 15, last annotation update)
 DE CDNA FLJ20197 F1S, CLONE COLF0996.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-COLON MUCOSA;
 RA Watanabe K., Kumagai A., Itakura S., Yamazaki M., Tashiro H., Ota T.,
 RA Suzuki Y., Ohtayashi M., Nishi T., Shibahara T., Tanaka T.,
 RA Nakamura Y., Isogai T., Sugano S.;
 RT "NEDO human cDNA sequencing project."
 RT Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AK000204; BAA91007.1;
 SQ SEQUENCE 782 AA; 88616 MW; F472FE54F627CE8 CRC64;

Query Match 11.6%; Score 103.5; DB 4; Length 782;
 Best Local Similarity 25.2%; Pred. No. 1.7;
 Matches 40; Conservative 32; Mismatches 48; Indels 39; Gaps 7;
 QY 54 RGLRAVME-----CRNTHLL---QOELTEAGKGFQDYEAQ-AAAC-----NHTVMA 97
 DB 284 RDSLHATAELQVRSVSLTHIALQEEELTRKVPSPDSLEPFRKQSLNMRREKVFA 343
 QY 98 LMASLDAEKAQKQKVEELEGETTTLNKH-----LQDASAEVERLRRENOV 143
 DB 344 LMVQLAKQLEHSDSVKQLKGVAISQEVYTSQSQEQAILQSLQDKAAEVEVERMGAKG 403
 QY 144 LSVRIA---ADKYPSSQDSSSAAPOLLIVLLGSA 177
 DB 404 LQELSLRAQEARRW---QQQTASAEQRLVYNAVSS 438

RESULT 8
 QY6XJ4 PRELIMINARY: PRT; 782 AA.
 AC QY6XJ4:
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, last annotation update)
 DE CDNA FLJ20210 F1S, CLONE COLF1787.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-COLON MUCOSA;
 RA Watanabe K., Kumagai A., Itakura S., Yamazaki M., Tashiro H., Ota T.,
 RA Suzuki Y., Ohtayashi M., Nishi T., Shibahara T., Tanaka T.,
 RA Nakamura Y., Isogai T., Sugano S.;
 RT "NEDO human cDNA sequencing project."
 RT Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL: AK000217; BAA91016.1; -
SQ SEQUENCE 782 AA; 88643 MW; A510C8BCFA8247B CRC64;

Query Match

Best Local Similarity 11.6%; Score 103.5; DB 4; Length 782;
Matches 40; Conservative 32; Mismatches 48; Indels 39; Gaps 7;

QY 54 RDGLRAVME-----CRNVTHLL---QOELTEAOKGFQDVEAQAATCNHTVMALMASIDEKAQOKRVE 97
DB 284 RDSLHAAVALLQVAVGSLTHLALQOELTEAOKGFQDVEAQAATCNHTVMALMASIDEKAQOKRVE 343
QY 98 LMASLDEKAQOKRVELESETTLNKK-----LQDASAEVERLRRENOV 143
DB 344 LMVQLKAQOELHSDSVKQGVASLQEKVTSQSEDAIILQRSIQDKRAAEVEVERMGAKG 403
QY 144 LSVRI-----ADKKYPPSSODSSAAPOLLIVLGLSA 177
DB 404 LQELSTRAOEARROW-----QOQTASAEQQLKLVYNAVSS 438

RESULT 9

028298

ID 028298 PRELIMINARY; PRT; 1534 AA.

AC 028298;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, last annotation update)
DE RIBOSOME RECEPTOR.
GN P180.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95310363; PubMed=7790375;
RA Wenker E.E., Sun Y., Savitz A.J., Meyer D.I.;
RT "Functional characterization of the 180 kDa ribosome receptor in
RT vivo."
RL J. Cell Biol. 130:29-39(1995).
DR EMBL: X87224; CA60676.1; -
KW Receptor.
SQ SEQUENCE 1534 AA; 164586 MW; B343BCF12656F3C5 CRC64;

Query Match

Best Local Similarity 11.5%; Score 102.5; DB 6; Length 1534;
Matches 26; Conservative 22; Mismatches 42; Indels 1; Gaps 1;

QY 55 DGLRAVMECRNVTHLLQOELTEAOKGFQDVEAQAATCNHTVMALMASIDEKAQOKRVE 114
DB 847 DAIVAKSKLEVENNELLAERAKAAGAKYKQLVAREQETLVAQANLEASVREHVEVQ 906
QY 115 ELEGITTLNKKLQDA-SAEVERLRRENOV 144
DB 907 QLOGKIRTLQOELNGPNTQARLQDENSIL 937

RESULT 10

096200

ID 096200 PRELIMINARY; PRT; 1627 AA.

AC 096200;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, last annotation update)
DE AXONEME-ASSOCIATED PROTEIN GASP-180.
OS Giardia lamblia (Giardia intestinalis).
OC Eukaryota; Diplomonadida; Hexamitidae; Giardinales; Giardia.
NCBI_TaxID=5741;
RN [1]
RP SEQUENCE FROM N.A.
RA Elendendorf H.G., Rohrer S.C., La Vigne E.A., Nash T.E.;

RT "Novel Axoneme-Associated Proteins in Giardia lamblia."
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF400249; AAK91740.1; -
SQ SEQUENCE 1627 AA; 179270 MW; 021EED9763907DCC CRC64;

Query Match

Best Local Similarity 11.5%; Score 102.5; DB 5; Length 1627;
Matches 34; Conservative 23; Mismatches 60; Indels 5; Gaps 3;

QY 46 IKANSEACROGLRAVMECRNVTHLLQOELTEAOKGFQDVEAQAATCNHTVMALMASL-DA 104
DB 1284 LRESAEALQDKLHLSORAAQDGLKVEQLERDLSGAKELVERATTIDELQRLRDT 1343
QY 105 EKAGOK-KVELEGETTLNKKLQDASAEVERLRRENOVSVRIADKKYPPSSODSSA 163
DB 1344 EBYDDLKRIALDELDELAVLNDGLKDDAEIAELREG---LEAOPATYVYPESGEVGD 1400
QY 164 AA 165
DB 1401 AA 1402

RESULT 11

091Y73

ID 091Y73 PRELIMINARY; PRT; 304 AA.

AC 091Y73;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, last annotation update)
DE MACROPHAGE GALACTOSE N-ACETYL-GALACTOSAMINE SPECIFIC LECTIN.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC014811; AHI14811.1; -
KW Lectin.
SQ SEQUENCE 304 AA; 34622 MW; 2272E1ADA2C0262A CRC64;

Query Match

Best Local Similarity 11.4%; Score 101.5; DB 11; Length 304;
Matches 36; Conservative 38; Mismatches 48; Indels 33; Gaps 7;

QY 24 LGIGILVLLIIVIGPLIIFTIKANSEACRD--GLRAVMECRNVTHLLQOELTEAOKGF 81
DB 41 LGSLILLVVVSVIG-----SONSQLRDRDGLTGLRAIID--NTSKIRAE-----F 83
QY 82 ODVAQAATCNHTVMALMASIDEKAQO-----KVELEGEI-----TTLNKKLQDAS 131
DB 84 QSLDSRADSEKGISSLKVDVHDHROLOAGRDLSQVTSLESTLEREALKTDLSDLT 143
QY 132 AEVERLRRENOVSVRIADKKYPPSSODSSAAP 166
DB 144 DHVQQLRKDLKALTCQLANLK---NNGSEVACCP 174

RESULT 12

09YHDS

ID 09YHDS PRELIMINARY; PRT; 826 AA.

AC 09YHDS;
DT 01-MAY-1999 (Tremblrel. 10, Created)
DT 01-MAY-1999 (Tremblrel. 10, last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, last annotation update)
DE MYOSIN HEAVY CHAIN (FRAGMENT).
HNC-4.
OS Rana catesbeiana (Bull frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranoidae; Ranidae; Rana.

QY	54	RDGJRAVNE-----CRNYTHLL-----QOEFLPAQKGFQDVENAQ--LATC-----NHPTMA	97
Db	8	RDSJHATAEILLQVNVQSTHTHTLALQEEELTRKQVPSDSLEPEFTRKQCSLLNRMREKVEFA	67
QY	98	LMASLDEPKAGQCKVETLEGEITTLNKK-----IQDASAE--VERLRKREN	14
Db	68	LMVQLKAELESHDSVSKQKGVASLQEKVTSQSQEQEALILORSLODKAAEIVGEBMGAKG	127
QY	142	QVLSVRIADKKRYPPSSQDSSAAAPQELLETVLLGLISA	177
Db	128	LQLELSIRAEQEARRRWQOOTAS--AEEOULFLVNNAVSS	162

RESULT	15	
P79391		
AC	P79391	PRELIMINARY; PRT; 270 AA.
DT	01-MAY-1997	(TREMBL) 03 Created
DT	01-MAY-1997	(TREMBL) 03 Last sequence update
DT	01-DEC-2001	(TREMBL) 19 Last annotation update
DE	LACTIN-LIKE OXIDIZED LDL RECEPTOR.	
OS	Bos taurus (Bovine).	
CC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Camelidae; Bovina; Caprinae.	

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

Qy

```

Query Match      100.0%  Score 889;  DB 1;  Length 180;
Best Local Similarity 100.0%  Pred. NO. 1.3e-63;
Matches 180;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

QY      1  MASTSYDCCRPMEGDKRKLILGLVLILVILVILGVPILFIITKANSEACRDLGRAV 60
|||||

```

DB 1 MASTSDYCVPMEDGDKRCKLLGICILVLLIIVLGVPLIIFTIKANSBACRDGIRAV 60
 QY 61 MECRNATHLLQOELTEAOKGFQDVEAQAATCNHTVMAIAMSLEDAEKAGCKVEELEGEI 120
 DB 61 MECRNATHLLQOELTEAOKGFQDVEAQAATCNHTVMAIAMSLEDAEKAGCKVEELEGEI 120
 QY 121 TITLNHKLQDASAEYERLRRENOYLSVRIADKKYTPSSQDSSAAAPOLLIVLGLSLAQ 180
 DB 121 TITLNHKLQDASAEYERLRRENOYLSVRIADKKYTPSSQDSSAAAPOLLIVLGLSLAQ 180

RESULT 2
 KUCR_RAT STANDARD; PRT; 550 AA.
 ID P10716;
 AC 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 KUPFER cell receptor.
 KCLR.
 Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 83-104.
 RX MEDLINE=88227939; PubMed=2836387;
 RA Hoyle G.W., Hill R.L.;
 RT "Molecular cloning and sequencing of a cDNA for a carbohydrate
 RT binding receptor unique to rat Kupfer cells.";
 RL J. Biol. Chem. 263:7487-7492(1988).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91107689; PubMed=1846367;
 RA Hoyle G.W., Hill R.L.;
 RT "Structure of the gene for a carbohydrate-binding receptor unique to
 RT rat Kupfer cells.";
 RL J. Biol. Chem. 266:1850-1857(1991).
 CC -1- FUNCTION: RECEPTOR WITH AN AFFINITY FOR GALACTOSE AND FUCOSE.
 CC COULD BE INVOLVED IN ENDOCYTOSIS.
 CC -1- SUBCELLULAR LOCATION: TYPE II membrane protein.
 CC -1- TISSUE SPECIFICITY: KUPFER CELLS.
 CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: J03734; AAA41472.1; -
 DR EMBL: M55532; AAA40892.1; -
 DR PIR: A28166; A28166.
 DR PIR: A38674; A38674.
 DR HSSP: P20693; 1HLJ.
 DR InterPro: IPR000017; Syntaxin.
 DR InterPro: IPR001304; lectin_c.
 DR Pfam: PF000059; lectin_c.1.
 DR SMART: SM00034; CLECT.1.
 DR SMART: SM00503; SYN.1.
 DR PROSITE: PS00615; C-TYPE_LLECTIN_1; 1.
 DR PROSITE: PS50041; C-TYPE_LLECTIN_2; 1.
 KW Receptor; Transmembrane; Glycoprotein; Lectin; Signal-anchor;
 KW Endocytosis.
 FT DOMAIN 1 42 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 43 69 SIGNAL-ANCHOR (TYPE-TI MEMBRANE PROTEIN)
 FT 70 550 (POTENTIAL).
 FT DOMAIN 438 538 EXTRACELLULAR (POTENTIAL).
 FT DISULFID 440 536 C-TYPE LECTIN (SHORT FORM).
 FT DISULFID 516 528 BY SIMILARITY.

FT CARBOHYD 87 87 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 93 93 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 115 115 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 132 132 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 209 209 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 255 255 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 550 AA; 61104 MW; 9358A6CF4C306270 CRC64;

Query Match 11.5%; Score 102; DB 1; Length 550;
 Best local Similarity 28.0%; Pred. No. 0.57; Mismatches 61; Indels 14; Gaps 3;
 Matches 37; Conservative 20;

QY 46 IKANSEACRDGLRAVMECRNVTHLLQOELTEAOKGFQDVEAQAATCNHTVMAIAMSLEDAE 105
 DB 256 ISAEIQAMRDGMQRAGE-----EMTSLKQLETLTAQIQNNNGHLEQDTQIQGL 305
 QY 106 KAQGCKVEELEGEITITLNHKLQDASAEYERLRRENOYLSVRIADKKYTPSSQDSSAAA 165
 DB 306 KAQ-LKSTSSLNSQIEVNGKLDKSSRELQTLRD---LSDVSALKSNVQMLQSLQAK 361
 QY 166 POLLIVLGLSA 177
 DB 362 AEVQSLKTGLEA 373

RESULT 3
 KICR_MOUSE STANDARD; PRT; 422 AA.
 ID P05784; Q61766;
 AC 01-NOV-1988 (Rel. 09, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Keratin, type I cytoskeletal 18 (Cytokeratin 18) (Cytokeratin endo B)
 DE (Keratin D).
 GN KRT18 OR KRT1-18 OR KERD.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89196920; PubMed=2467843;
 RA Ichihose Y., Morita T., Zhang F., Srinahesonggram S., Tondella M.L.C.,
 RA Matsumoto M., Nozaki M., Matsushiro A.;
 RT "Nucleotide sequence and structure of the mouse cytokeatin endob
 RT gene.";
 RL Gene 70:85-95(1988).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Teratocarcinoma;
 RA Alonso A., Weber T., Jorcano J.L.;
 RT "Cloning and characterization of keratin D, a murine endodermal
 RT cytoskeletal protein induced during in vitro differentiation of F9
 RT teratocarcinoma cells.";
 RL Roux's Arch. Dev. Biol. 196:16-21(1987).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86085876; PubMed=2416755;
 RA Singer P.A., Trevor K., Oshima R.G.;
 RT "Molecular cloning and characterization of the Endo B cytokeatin
 RT expressed in preimplantation mouse embryos.";
 RL J. Biol. Chem. 261:538-547(1986).
 RN [4]
 RP SEQUENCE OF 1-131 FROM N.A.
 RX MEDLINE=88255838; PubMed=2454868;
 RA Oshima R.G., Trevor K., Shevinsky L.H., Ryder O.A., Cecena G.;
 RT "Identification of the gene coding for the Endo B murine cytokeatin
 RT and its methylated, stable inactive state in mouse nonepithelial
 RT cells.";
 RL Genes Dev. 2:505-516(1988).
 CC -1- SUBUNIT: HETEROPTRAMER OF TWO TYPE I AND TWO TYPE II KERATINS.
 CC KERATIN 18 ASSOCIATES WITH KERATIN 8.

```

CC -1- MISCELLANEOUS: THERE ARE TWO TYPES OF CYTOSKELETAL AND
CC MICROFIBRILLAR KERATIN, I (ACIDIC) AND II (NEUTRAL TO BASIC)
CC (40-55 AND 56-70 KILODALTONS, RESPECTIVELY).
CC -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M2832; AAA37552.1; -
DR EMBL: M36376; AAA39373.1; -
DR EMBL: M11686; AAA39390.1; -
DR EMBL: Y00217; CAA68365.1; -
DR PIR: A25621; A25621.
DR PIR: A28428; A28428.
DR PIR: J70406; J70406.
DR SWISS-2DPAGE: P05784; MOUSE.
DR MGI: 96692; Krt1-18.
DR InterPro: IPR001664; IF.
DR InterPro: IPR002957; Keratin_1.
DR Pfam: PF00038; filament_1.
DR PRINTS: PR01248; TYPE1KERATIN.
DR PROSITE: PS00226; IF: 1.
KW Intermediate filament; Coiled coil; Keratin; Glycoprotein;
KW Acetylation.
FT INT_MET 0 0 BY SIMILARITY.
FT MOD_RES 1 1 ACETYLATION (BY SIMILARITY).
FT DOMAIN 1 70 HEAD.
FT DOMAIN 380 422 ROD.
FT DOMAIN 71 379 TAIL.
FT DOMAIN 107 106 COIL 1A.
FT DOMAIN 125 124 COIL 1B.
FT DOMAIN 217 240 LINKER 1.
FT DOMAIN 241 379 LINKER 12.
FT SITE 263 263 COIL 2.
FT SITE 323 323 STUTTER.
FT CARBOHYD 30 30 O-LINKED (GLCNAC) (BY SIMILARITY).
FT CARBOHYD 31 31 O-LINKED (GLCNAC) (BY SIMILARITY).
FT CARBOHYD 49 49 O-LINKED (GLCNAC) (BY SIMILARITY).
FT CONFLICT 133 133 L -> F (IN REF. 1).
FT CONFLICT 243 243 D -> N (IN REF. 2).
FT CONFLICT 252 252 A -> G (IN REF. 2).
SQ SEQUENCE 422 AA; 47373 MW; 4D5B0E9C7732F2F3 CRC64;

```

```

ID MGL_MOUSE STANDARD: PRT; 304 AA.
AC P49300;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Macrophage asialoglycoprotein-binding protein (M-ASGP-BP) (Macrophage
DE galactose/N-acetylgalactosamine-specific lectin) (MGL).
GN MGL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C3H/HEN;
RX MEDLINE=92268032; PubMed=1587794;
RA Sato M., Kawakami K., Osawa T., Toyoshima S.;
RT "Molecular cloning and expression of cDNA encoding a galactose/N-
RT acetylgalactosamine-specific lectin on mouse tumoricidal
RT macrophages."
RL J. Biochem. 111:331-336(1992).
RN [2]
RP SEQUENCE OF 102-120 AND 137-151.
RC STRAIN=C3H/HEN;
RX MEDLINE=89197865; PubMed=3241002;
RA Oda S., Sato M., Toyoshima S., Osawa T.;
RT "Purification and characterization of a lectin-like molecule specific
RT for galactose/N-acetyl-galactosamine from tumoricidal macrophages."
RL J. Biochem. 104:600-605(1988).
CC -1- FUNCTION: RECOGNIZES TERMINAL GALACTOSE AND N-ACETYLGALACTOSAMINE
CC UNITS. MAY PARTICIPATE IN THE INTERACTION BETWEEN TUMORICIDAL
CC MACROPHAGES AND TUMOR CELLS.
CC -1- SUBUNIT: HOMO-OLIGOMER.
CC -1- SUBCELLULAR LOCATION: Type II membrane protein.
CC -1- TISSUE SPECIFICITY: IS EXPRESSED ON THE SURFACE OF ACTIVATED
CC MACROPHAGES.
CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: S36676; AB22171.1; -
DR HSSP: P06734; IKCE.
DR MGI: 96975; Mgl.
DR InterPro: IPR001304; lectin_c.
DR Pfam: PF00059; lectin_c; 1.
DR SMART: SM00034; CLECT; 1.
DR PROSITE: PS00615; C-TYPE_LLECTIN_1; 1.
DR PROSITE: PS00615; C-TYPE_LLECTIN_2; 1.
KW Lectin; Glycoprotein; Transmembrane; Calcium; Signal-anchor.
FT DOMAIN 1 35 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 36 56 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT DOMAIN 57 304 (POTENTIAL).
FT DOMAIN 172 298 EXTRACELLULAR (POTENTIAL).
FT DISULFID 173 184 C-TYPE LECTIN (LONG FORM).
FT DISULFID 201 296 BY SIMILARITY.
FT DISULFID 274 288 BY SIMILARITY.
FT CARBOHYD 74 74 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 166 166 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 304 AA; 34596 MW; 3F79CD12C34F5BCC CRC64;

```

```

Query Match 11.4%; Score 101; DB 1; Length 422;
Best Local Similarity 27.4%; Pred. No. 0.51; Mismatches 56; Indels 50; Gaps 8;
Matches 51; Conservative 27; Mismatches 56; Indels 50; Gaps 8;

```

```

Query Match 11.2%; Score 99.5; DB 1; Length 304;
Best Local Similarity 23.2%; Pred. No. 0.47; Mismatches 36; Indels 33; Gaps 7;
Matches 36; Conservative 37; Mismatches 49; Indels 33; Gaps 7;

```

```

Db      41 LGISLLLVVSVIG-----SONSLRDLGTLFATLD--NTTSKIAE-----F 83
Oy      82 ODEVAOATCNHTVMALMSLDAEKAQO-----KKVEELGEGI-----TTLNKKILODAS 131
Db      84 QSDSRADSEKSISSKUYVEDHROELQGRDLSQKRVTSLESTVEKREBALKTDLSDLT 143
Oy      132 AEVERLRRENOVLVRADKKRYPPSSODSSSAAP 166
Db      144 DHVOQLRKDLKALTCOLANLK-----NNGSEVACCP 174

RESULT 5
VDP_HUMAN
ID      VDP_HUMAN          STANDARD;          PRT;          962 AA.
AC      060763;
DT      16-OCT-2001 (Rel. 40, Last sequence update)
DT      16-OCT-2001 (Rel. 40, Last annotation update)
DE      General vesicular transport factor p115 (Transcytosis associated protein) (TAP) (Vesicle docking protein).
GN      VDP.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OX      Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
RN      NCBI_TaxId=9606;
RP      SEQUENCE FROM N.A., AND PHOSPHORYLATION SITE SER-942.
RX      MEDLINE-98148093; PubMed-9478999;
RT      "Phosphorylation of the vesicle docking protein p115 regulates its association with the Golgi membrane."
RL      J. Biol. Chem. 273:5385-5388(1998).
CC      -1- FUNCTION: GENERAL VESICULAR TRANSPORT FACTOR REQUIRED FOR INTERCYTOSOLIC FUSION AND/OR SUBSEQUENT BINDING OF THE VESICLES TO THE TARGET MEMBRANE. MAY WELL ACT AS A VESICULAR ANCHOR BY INTERACTING WITH THE TARGET MEMBRANE AND HOLDING THE VESICULAR AND TARGET MEMBRANES IN PROXIMITY (BY SIMILARITY).
CC      -1- SUBCELLULAR LOCATION: PERIPHERAL MEMBRANE PROTEIN WHICH RECYCLES BETWEEN THE CYTOSOL AND THE GOLGI APPARATUS DURING INTERPHASE.
CC      -1- DOMAIN: COMPOSED OF A GLOBULAR HEAD, AN ELONGATED TAIL (COILED-COIL) AND A HIGHLY ACIDIC C-TERMINAL DOMAIN.
CC      -1- PTM: PHOSPHORYLATED IN A CELL CYCLE-SPECIFIC MANNER; PHOSPHORYLATED IN INTERPHASE BUT NOT IN MITOTIC CELLS.
CC      DEPHOSPHORYLATED PROTEIN ASSOCIATES WITH THE GOLGI MEMBRANE; PHOSPHORYLATION PROMOTES DISSOCIATION.
CC      -1- SIMILARITY: BELONGS TO THE VDP/USOL/YB1047C FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL, D86326; BAA25300.1; -.
CC      DR      MIM: 603344; -.
CC      DR      InterPro: IPR000225; Armadillo.
CC      DR      PROSITE: PS50176; ARM_REPEAT; UNKNOWN_1.
CC      KW      Transport; Protein transport; Golgi stack; Membrane; Coiled coil; Phosphorylation.
CC      FT      DOMAIN 1 637 GLOBULAR HEAD.
CC      FT      DOMAIN 638 930 COILED COIL (POTENTIAL).
CC      FT      DOMAIN 935 962 ASP/GLU-RICH (ACIDIC).
CC      FT      MOD_RES 942 942 PHOSPHORYLATION.
CC      FT      MUTAGEN 942 942 S->A: LOSS OF PHOSPHORYLATION.
CC      SQ      SEQUENCE 962 AA; 107906 MM; 2E748E2C1BCB942 CRC64;

```

```

Matches 35; Conservative 26; Mismatches 54; Indels 8; Gaps 3;
Oy      60 VMECRNVTHTLLOELPEAGQFQDVE-AQAATCNHTVMALMSLDAEKAQOKKVELEG 118
Db      739 IEELKRNOELLQSLTEKDSMTENKSSQTSNGNESSATVSRDSE-----QVANELKO 792
Oy      119 EITTLNKKILODASAEVERLRRENO-VLSVRIADKKRYPPSSODSSSAAPOLLIVLGLSA 177
Db      793 ELATLKSQLSQSVETTKLOTEKREKQELLQTEAKSAKSVENGQETITLARTTVEGRSLA 852
Oy      178 LLQ 180
Db      853 LLQ 855

RESULT 6
MYSB_CAEEL
ID      MYSB_CAEEL          STANDARD;          PRT;          1966 AA.
AC      P02566;
DT      21-JUL-1986 (Rel. 01, Created)
DT      21-JUL-1986 (Rel. 01, Last sequence update)
DT      16-OCT-2001 (Rel. 40, Last annotation update)
DE      Myosin heavy chain B (MHC B).
GN      UNC-54 OR MYO-4.
OS      Caenorhabditis elegans.
OC      Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae;
OX      Rhabditidae; Peloderinae; Caenorhabditis.
RN      NCBI_TaxId=6239;
RP      SEQUENCE FROM N.A.
RX      MEDLINE-83273600; PubMed-6576334;
RA      Karn J., Brenner S., Barnett L.;
RT      "Protein structural domains in the Caenorhabditis elegans unc-54 myosin heavy chain gene are not separated by introns."
RL      Proc. Natl. Acad. Sci. U.S.A. 80:4253-4257(1983).
RN      [12]
RP      SEQUENCE OF 850-1966 FROM N.A.
RX      MEDLINE-82272395; PubMed-7202124;
RA      McLachlan A.D., Karn J.;
RT      "Periodic charge distributions in the myosin rod amino acid sequence RT match cross-bridge spacings in muscle."
RL      Nature 299:226-231(1982).
RN      [13]
RP      SEQUENCE OF 1876-1966 FROM N.A.
RX      MEDLINE-83232892; PubMed-6571695;
RA      Wills N., Gesteland R.F., Karn J., Barnett L., Bolten S., Waterston R.H.;
RT      "The genes sup-7 x and sup-5 III of C. elegans suppress amber nonsense mutations via altered transfer RNA."
RL      Cell 33:575-583(1983).
CC      -1- FUNCTION: MUSCLE CONTRACTION.
CC      -1- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2 HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC) AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
CC      -1- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
CC      -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES, CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
CC      -1- PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.
CC      -1- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT MEROMYOSIN (LMM) AND 1 HEAVY MEROMYOSIN (HMM). IT CAN LATER BE SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED SUBFRAGMENT (S2).
CC      -1- MISCELLANEOUS: THERE ARE FOUR DIFFERENT MYOSIN HEAVY CHAINS IN C.ELEGANS.
CC      -1- MISCELLANEOUS: MHC A AND MHC B ARE FOUND EXCLUSIVELY IN THE BODY WALL MUSCLE. THEY CO-ASSEMBLE INTO BODY WALL THICK FILAMENT.
CC      -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its

```

Query Match 11.1%; Score 99; DB 1; Length 962;
Best Local Similarity 28.5%; Pred. No. 1.8;

CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: J01050; AAA28124.1; -
 DR EMBL: V01494; CAA24738.1; -
 DR PIR: A02992; MMKW.
 DR HSSP: P08799; LMND.
 DR InterPro: IPR004009; Myosin_N.
 DR InterPro: IPR002928; Myosin_tail.
 DR InterPro: IPR001609; myosin_head.
 DR Pfam: PF00063; myosin_head.1.
 DR Pfam: PF02736; Myosin_N.1.
 DR Pfam: PF01576; Myosin_tail.1.
 DR PRINTS: PR00193; MYOSINHEAVY.
 DR ProDom: PD000355; myosin_head; 1.
 DR SMART: SM00242; MYSC; 1.
 KW Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;
 KM ATP-binding; Methylation; Alkylation; Multigene family.
 FT DOMAIN 1 850 MYOSIN HEAD-LIKE.
 FT DOMAIN 851 1966 COILED COIL (POTENTIAL).
 FT DOMAIN 1165 1176 ALPHA-HELICAL TAILPIECE (S2).
 FT DOMAIN 1165 1966 LIGHT MEROMYOSIN (LMM).
 FT NP_BIND 177 184 ATP (BY SIMILARITY).
 FT DOMAIN 665 687 ACTIN-BINDING.
 FT DOMAIN 769 783 ACTIN-BINDING.
 FT MOD_RES 128 128 METHYLATION (TR1-) (POTENTIAL).
 FT MOD_RES 705 705 ALKYLATION (SH-1).
 FT MOD_RES 715 715 ALKYLATION (SH-2).
 FT CONFLICT 1337 1337 E -> R (IN REF. 2).
 FT CONFLICT 1880 1880 I -> L (IN REF. 2).
 FT SEQUENCE 1966 AA; 225125 MW; B6F0BB2FE27B67F CRC64;

Query Match 11.1%; Score 98.5; DB 1; Length 1966;
 Best Local Similarity 24.7%; Pred. No. 4.2; Mismatches 56; Indels 31; Gaps 5;
 Matches 36; Conservative 21;

QY 47 KANSEACRDGLRAVMECRNVTHLLOQELTEAOKFPDVEAQAATCNHTVMALMASIDAEK 106
 DB 1811 EEAFAALGKGKRVIAKLEQVRVLESELDGEORFODANKNGRADRVRELOFQVDEDK 1870
 QY 107 AQ-----GOKKVELEGEITTLN-----HKQDASAEVERLRREN 141
 DB 1871 KNEERLQDLIDKLOOKLKTKQKOVEAE--ELANLNLOKYKOLTHOLEDAEERAD--QAEN 1927
 QY 142 QVLAVRI--ADKRYPPSSODSSAA 164
 DB 1928 SLSKMRSKRASAVAPGLQSSASAA 1953

RESULT 7
 MYH4_RABIT STANDARD; PRT; 1938 AA.
 AC 028641:
 DT 16-OCT-2001 (Rel. 40; Created)
 DT 16-OCT-2001 (Rel. 40; Last sequence update)
 DT 16-OCT-2001 (Rel. 40; Last annotation update)
 DE Myosin heavy chain, skeletal muscle, juvenile.
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 OX NCBI_TaxID=9986;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NEW ZEALAND WHITE; TISSUE=Skeletal muscle;
 RA Maeda K., Hostinova E., Roesch-Kleinkauf A., Schuster H., Gasperik J.,
 Wittinghofer A.;
 RT "Isolation, sequencing of myosin heavy chain cDNA from rabbit
 skeletal muscle and a novel cosynthesis of S-1 fragment with the
 essential and regulatory light chains.";

RL Submitted (AUG-1995) to the EMBL/Genbank/DBJ databases.
 CC -1- SUBUNIT: MUSCLE CONTRACTION.
 CC -1- FUNCTION: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
 CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)
 CC AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
 CC -1- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
 CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPEITIVE, SHOWING
 CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
 CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
 CC -1- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
 CC MEROMYOSIN (LMM) AND 1 HEAVY MEROMYOSIN (HMM). IT CAN LATTER BE
 CC SPLIT FURTHER INTO 2 GLOBULAR SUPERFRACTIONS (S1) AND 1 ROD-SHAPED
 CC SUBFRACTION (S2).
 CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 IQ DOMAIN.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: U32574; AAA74199.1; -
 DR HSSP: P08799; LMND.
 DR InterPro: IPR000048; IQ.
 DR InterPro: IPR004009; Myosin_N.
 DR InterPro: IPR002928; Myosin_tail.
 DR InterPro: IPR001609; myosin_head.
 DR Pfam: PF00063; myosin_head.1.
 DR Pfam: PF02736; Myosin_N.1.
 DR Pfam: PF01576; Myosin_tail.1.
 DR PRINTS: PR00193; MYOSINHEAVY.
 DR ProDom: PD000355; myosin_head; 1.
 DR SMART: SM00015; IQ; 1.
 DR SMART: SM00242; MYSC; 1.
 DR PROSITE: PS50096; IQ; 1.
 KW Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;
 KM Calmodulin-binding; ATP-binding; Methylation; Alkylation;
 KM Multigene family.
 FT DOMAIN 1 783 MYOSIN HEAD-LIKE.
 FT DOMAIN 784 813 IQ.
 FT NP_BIND 179 186 ATP (POTENTIAL).
 FT DOMAIN 658 680 ACTIN-BINDING (BY SIMILARITY).
 FT DOMAIN 774 774 ACTIN-BINDING (BY SIMILARITY).
 FT MOD_RES 35 35 METHYLATION (MONO-) (BY SIMILARITY).
 FT MOD_RES 130 130 METHYLATION (TR1-) (BY SIMILARITY).
 FT MOD_RES 552 552 METHYLATION (TR1-) (BY SIMILARITY).
 FT MOD_RES 756 756 METHYLATION (TR1-) (BY SIMILARITY).
 FT MOD_RES 698 698 ALKYLATION (SH-1) (BY SIMILARITY).
 FT MOD_RES 708 708 ALKYLATION (SH-2) (BY SIMILARITY).
 FT SEQUENCE 1938 AA; 223064 MW; D8A8A2EC5B182626 CRC64;

Query Match 11.0%; Score 98; DB 1; Length 1938;
 Best Local Similarity 26.1%; Pred. No. 4.6; Mismatches 45; Indels 40; Gaps 5;
 Matches 37; Conservative 20;

QY 47 KANSEACR-----DGLRAVMECRNVTHLLOQELTEAOKFPDVEAQAATCNHTVMALM 99
 DB 1366 KANSEVAQMKTKYETAIDQTELEAKKKLAIRDAD--EHEAVNKKC----- 1414
 QY 100 ASDAERAGOKKVELEGEITTLN-----HKQDASAEVERLR 138
 DB 1415 ASLEKTKRQNRQNVDELMDIVERTNAACAALDKKQNPDKILAEKMKHKEETFALEASQ 1474
 QY 139 RENQVLSVRIAD-KKRYPPSSOD 159
 DB 1475 KESRSLSSTVEFKYKNAYEESLD 1496

RESULT 8
MYSD_CAEEL STANDARD: PRT: 1938 AA.
ID MYSD_CAEEL 019674: PRT: 1938 AA.
AC P02567: 019674: PRT: 1938 AA.
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Myosin heavy chain D (MHC D).
GN MYO-1 OR R06C7.10.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peleodermata; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX MEDLINE=89178677; PubMed=2926820;
RA Dobb N.J., Maruyama I.N., Krause M., Karn J.;
RT "Sequence analysis of the complete Caenorhabditis elegans myosin
heavy chain gene family.";
RL J. Mol. Biol. 205:603-613(1989).
RN [2]
RP SEQUENCE OF 34-1795 FROM N.A.
RX MEDLINE=83273600; PubMed=6576334;
RA Karn J., Brenner S., Barnett L.;
RT "Protein structural domains in the Caenorhabditis elegans unc-54
myosin heavy chain gene are not separated by introns.";
RL Proc. Natl. Acad. Sci. U.S.A. 80:4253-4257(1983).
RN [3]
RP SEQUENCE OF 115-365 AND 1492-1763 FROM N.A.
RX MEDLINE=85201409; PubMed=3888374;
RA Karn J., Dobb N.J., Miller D.M.;
RT "Cloning nematode myosin genes.";
RL Cell Muscle Motil. 6:185-237(1985).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Gardner A., McMurray A.;
RT Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
RN [5]
RP FUNCTION: MUSCLE CONTRACTION.
CC -1- SUBUNIT: MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)
AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
CC -1- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
CC -1- TISSUE SPECIFICITY: FOUND EXCLUSIVELY IN THE PHARYNGEAL MUSCLE.
CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
CC -1- PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY
ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.
CC -1- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
MEROMYOSIN (LMW) AND 1 HEAVY MEROMYOSIN (HMM). IT CAN LATER BE
SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
SUBFRAGMENT (S2).
CC -1- MISCELLANEOUS: THERE ARE FOUR DIFFERENT MYOSIN HEAVY CHAINS IN
C.ELEGANS.
CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X08065; CAA30854.1; -
DR EMBL: M37232; AAA28119.1; -
DR EMBL: M37234; AAA28120.1; -
DR EMBL: 271266; CAA95848.1; -
DR EMBL: 271261; CAA95848.1; JOINED.
DR EMBL: 271261; CAA95806.1; -

DR EMBL: 271266; CAA95806.1; JOINED.
DR PIR: S02772; MKMK1.
DR HSSP: P08799; 1MND.
DR WormRep: R06C7.10; CE06253.
DR InterPro: IPR004009; Myosin_N.
DR InterPro: IPR002928; Myosin_tail.
DR InterPro: IPR001609; myosin_head.
DR Pfam: PF00063; myosin_head.
DR Pfam: PF02736; myosin_N; 1.
DR Pfam: PF01576; Myosin_tail; 1.
DR PRINTS: PRO0193; MYOSINHEAVY.
DR ProDom: PD000355; myosin_head; 1.
DR SMART: SM00242; MYSC; 1.
DR ATP-binding: Methylation; Coiled coil; Thick filament; Actin-binding;
KW ATP-binding; Methylation; Alkylation; Multigene family.
FT DOMAIN 1 845
FT MOD.RES 128 128
FT MOD.RES 700 700
FT MOD.RES 710 710
FT MOD.RES 94 94
FT CONFLICT 98 98
FT CONFLICT 377 377
FT CONFLICT 389 390
FT CONFLICT 391 391
FT CONFLICT 408 408
FT CONFLICT 474 474
FT CONFLICT 577 577
FT CONFLICT 681 681
FT CONFLICT 1373 1373
FT CONFLICT 1659 1659
SQ SEQUENCE 1938 AA; 223255 MW; 387399C8F634ACF4 CRC64;
Query Match 11.0%; Score 98; DB 1; Length 1938;
Best Local Similarity 26.6%; Pred. No. 4.6;
Matches 34; Conservative 24; Mismatches 58; Indels 12; Gaps 4;
QY 51 EACRDGIRAME-----CRVTHLLOQ--ELTEAO-KGFQDVEAQAATCNHTVMAIMASLD 103
DB 1324 KANDELHEHGEHFHACKNLEHEDQCHELLEEDQINKDIOQLSINSIEISOMKARYE 1383
QY 104 AEKAGQCKVVELEGETITTNHKLQDASAEVERLRRENOVLSVRIADKTYPSQDSSSA 163
DB 1384 GEGIVGSEEELEELKRNKMNREVMDOELMSAA-----ONKVISLEKAKGKLAEETEDARSD 1438
QY 164 AAPQLLIIV 171
DB 1439 VDRHLYIV 1446
RESULT 9
PU92_SCICO STANDARD: PRT: 286 AA.
AC P22312;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE PUF1.11/9-2 protein precursor.
GN 11/9-2.
OS Sciaridae coprophila (Fungus gnaf).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Sciaridae;
OC Sciaridae; Brachyura.
OX NCBI_TaxID=38358;
RN [1]
RP SEQUENCE FROM N.A.

DE Hypothetical 100.0 kDa protein M01A8.2 in chromosome III.
 GN M01A8.2.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN
 RP
 RC SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RX MEDLINE=94150718; PubMed=7906398;
 RA Wilson R., Almscough R., Anderson K., Baynes C., Berks M.,
 RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
 RA Craxton M., Dear S., Du Z., Durdin R., Favell A., Fraser A.,
 RA Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
 RA Johnston L., Jones M., Kershaw J., Kirsten J., Laister N.,
 RA Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
 RA Parsons J., Percy C., Ritken L., Roopra A., Saunders D., Shownkeen R.,
 RA Sims M., Smalodon N., Smith A., Smith M., Sonhammer E., Staden R.,
 RA Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
 RA Waterston J., Watson A., Weinstock L., Wilkinson-Sproat J.,
 RA Wohldman P.;
 RT *2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 RT elegans.";
 RL Nature 368:32-38(1994).
 CC -1- SIMILARITY: CONTAINS 1 CAP-GLY DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: 227081; CAA81607.1; -
 DR PIR: S40998; S40998.
 DR PIR: S40999; S40999.
 DR WormPep: M01A8.2; GE03491.
 DR InterPro: IPR000938; CAP-GLY.
 DR Pfam: PF01302; CAP-GLY.1.
 DR PROSITE: PS00845; CAP-GLY.1; 1.
 DR PROSITE: PS50245; CAP-GLY.2; 1.
 KW Hypothetical protein; Coiled coil.
 FT DOMAIN 39 81 CAP-GLY.
 FT DOMAIN 522 696 COILED COIL (POTENTIAL).
 FT DOMAIN 729 756 COILED COIL (POTENTIAL).
 SQ SEQUENCE 893 AA; 99997 MW; 464F2962B36C28B1 CRC64;

Query Match 10.9%; Score 96.5; DB 1; Length 893;
 Best Local Similarity 24.8%; Pred. No. 2.6;
 Matches 32; Conservative 29; Mismatches 39; Indels 29; Gaps 4;

QY 48 ANSEARQDGLRAMECRNVTYHLLD--OELTEAOKGFQDV---EAQAATGNHTYMALMAS 101
 DB 566 SNOQVYRNHANNV-ESIQKTHETQIAEKKEFEERREARRDAEVCAMNRRHQKVAC 624
 QY 102 LDAEKAQKQKVELE-----GEITLNNHKLQDASAEVERLR 139
 DB 625 LDKIEAEKQCEQDLNKKVLAALANDCDHNNOMLTRKISSQTLALKKSAEMKELRQ 684
 QY 140 ENQVLSVRI 148
 DB 685 KNQNLISQV 693

RESULT 12
 PU91_SCICO STANDARD; PRT; 286 AA.
 AC P22311;
 DT 01-AUG-1991 (Rel. 19, Created)
 DT 01-AUG-1991 (Rel. 19, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE PUF II/9-1 protein precursor.
 GN II/9-1.
 OS Sciara coprophila (Fungus gnat).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Sciarioidea;
 OC Sciariidae; Brachyia.
 OX NCBI_TaxID=38358;
 RN
 RP
 RC SEQUENCE FROM N.A.
 RC STRAIN=6980;
 RX MEDLINE=90133907; PubMed=2614832;
 RA Dibartolomeis S.M., Gerbl S.A.;
 RA "Molecular characterization of DNA puf II/9A genes in Sciara
 RT coprophila.";
 RL J. Mol. Biol. 210:531-540(1989).
 CC -1- MISCELLANEOUS: THE DNA PUF II/9 PROTEINS HAVE A PROPOSED
 CC INTERMOLECULAR COILED COIL STRUCTURE WITH POSSIBLY INTERMOLECULAR
 CC DISULFIDE BRIDGES FORMED BY NUMEROUS CYSTEINE RESIDUES IN POSITION
 CC D OF THE HEPTAD REPEAT.
 CC -1- SIMILARITY: 768 IDENTICAL TO THE PUF II/9-2 PROTEIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X51680; CAA35983.1; -
 DR PIR: S07532; S07532.
 DR Signal; Coiled coil; Glycoprotein.
 KW Signal; Coiled coil; Glycoprotein.
 FT SIGNAL 1 19 OR 21 (POTENTIAL).
 FT CHAIN 20 286 PUF II/9-1 PROTEIN.
 FT DOMAIN 61 235 HELICAL (POTENTIAL).
 FT CARBOHYD 156 156 N-LINKED (GLCNAC...) (POTENTIAL).
 SQ SEQUENCE 286 AA; 32034 MW; AA6A7B55F191BB1D CRC64;

Query Match 10.8%; Score 96; DB 1; Length 286;
 Best Local Similarity 30.0%; Pred. No. 0.84;
 Matches 24; Conservative 16; Mismatches 36; Indels 4; Gaps 1;

QY 59 AWEBCRNVTYHLLDQELTEAOKGFQDVEAQAATGNHTYMALMASLDAEKAQKQKVELEG 118
 DB 102 ALCECQNSNLSLQTIQTLKELAQTKQELANCKEA-----LANCKAEMARKLKKIEELNC 157
 QY 119 EITLNNHKLQDASAEVERLR 138
 DB 158 TITQLEDEQCRARRERDQ 177

RESULT 13
 MYH8_HUMAN STANDARD; PRT; 1937 AA.
 ID MYH8_HUMAN
 AC P13535; Q14910;
 DT 01-JUN-1990 (Rel. 13, Created)
 DT 15-JUN-1998 (Rel. 36, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Myosin heavy chain, skeletal muscle, perinatal (MyHC-perinatal).
 GN MYH8.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN
 RP
 RC SEQUENCE FROM N.A.
 RC TISSUE=Skeletal muscle;
 RC MEDLINE=9033631; PubMed=2373371;
 RA Karsch-Mizirachi I., Peghali R., J. B. J., Leitwand L.A.;
 RT "Generation of a full-length human perinatal myosin heavy-chain-
 RL encoding cDNA.";
 Gene 89:289-294(1990).

[2]
RP SEQUENCE FROM N.A.
RC TISSUE-Skeletal muscle;
RX MEDLINE=95324556; PubMed=7601129;
RA Julian E.H., Kelly A.M., Pompidou A.J., Hoffman R., Schiaffino S.,
RT Steadman H.H., Rubinstein N.A.,
RT "Characterization of a human perinatal myosin heavy-chain
transcript.";
RL Eur. J. Biochem. 230:1001-1006(1995).
[3]
RP SEQUENCE OF 502-1937 FROM N.A.
RC TISSUE-Skeletal muscle;
RX MEDLINE=90235862; PubMed=1691980;
RA Bober E., Buchberger-Seidl A., Braun T., Singh S., Goedde H.W.,
RT Arnold H.H.,
RT "Identification of three developmentally controlled isoforms of human
myosin heavy chains.";
RL Eur. J. Biochem. 189:55-65(1990).
[4]
RP SEQUENCE OF 860-1937 FROM N.A.
RX MEDLINE=89234168; PubMed=2715179;
RA Feghali R., Leinwand L.A.,
RT "Molecular genetic characterization of a developmentally regulated
human perinatal myosin heavy chain.";
RL J. Cell Biol. 108:1791-1797(1989).
[5]
RP SEQUENCE OF 1-46 FROM N.A.
RA Esser K., Tidhar A., Myszkowski M.,
RT "Isolation and characterization of the human perinatal MHC promoter.";
RL Submitted (May-1998) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: MUSCLE CONTRACTION.
CC -1- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)
AND 2 REGULATOR LIGHT CHAIN SUBUNITS (MLC-2).
CC -1- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
CC -1- PFM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY
ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.
CC -1- MEROMYOSIN (LM) AND 1 HEAVY MEROMYOSIN (HMM). IT CAN LATER BE
SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
SUBFRAGMENT (S2).
CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 IQ DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@sib.ch).

DR EMBL: M36769; AAC17185.1; -
DR EMBL: Z38133; CA86293.1; -
DR EMBL: X51592; CA35841.1; -
DR EMBL: M35250; AA36546.1; -
DR EMBL: AF067143; AAC21557.1; -
DR PIR: A30220; A30220.
DR HSSP: P08799; ILVK.
DR MIM: 160741; -
DR InterPro: IPR000048; IQ.
DR InterPro: IPR004009; Myosin_N.
DR InterPro: IPR002928; Myosin_tail.
DR InterPro: IPR001609; myosin_head.
DR Pfam: PF00612; IQ; 1.
DR Pfam: PF00063; myosin_head; 1.
DR Pfam: PF02736; myosin_N; 1.
DR Pfam: PF01576; myosin_tail; 1.
DR PRINTS: PR00193; MYOSINHEAVY.
DR ProDom: PD000355; myosin_head; 1.

DR SMART: SM00015; IQ; 1.
DR SMART: SM00242; MYSC; 1.
DR PROSITE: PS50096; IQ; 1.
KW Myosin, muscle protein; Coiled coil; Thick filament; Actin-binding;
KW ATP-binding; Methylation; Alkylation; Multigene family;
KW Calmodulin-binding.
FT DOMAIN 1 780
FT MYOSIN HEAD-LIKE.
FT DOMAIN 81 813
FT COILED COIL (POTENTIAL).
FT NP_BIND 842 1937
FT ATP.
FT DOMAIN 181 188
FT ACTIN-BINDING.
FT DOMAIN 658 680
FT ACTIN-BINDING.
FT MOD_RES 132 132
FT METHYLATION (SH-1) (POTENTIAL).
FT MOD_RES 698 698
FT ALKYLATION (SH-2) (POTENTIAL).
FT MOD_RES 708 708
FT A -> R (IN REF. 2).
FT CONFLICT 15 15
FT E -> Q (IN REF. 1 AND 4).
FT CONFLICT 970 970
FT M -> N (IN REF. 3).
FT CONFLICT 1072 1072
FT N -> H (IN REF. 1 AND 4).
FT CONFLICT 1247 1247
FT MC -> DSG (IN REF. 3).
FT CONFLICT 1251 1252
FT E -> G (IN REF. 1 AND 4).
FT CONFLICT 1261 1261
FT K -> Q (IN REF. 1 AND 4).
FT CONFLICT 1297 1297
FT KY -> NT (IN REF. 3).
FT CONFLICT 1377 1378
FT EN -> AH (IN REF. 1 AND 4).
FT CONFLICT 1504 1505
FT E -> D (IN REF. 1 AND 4).
FT CONFLICT 1847 1847
FT D -> H (IN REF. 2).
FT CONFLICT 1914 1914
SQ SEQUENCE 1937 AA; 222762 MW; A3EE2D151792E9E8 CRC64;

Query Match 10.7%; Score 95; DB 1; Length 1937;
Best Local Similarity 25.2%; Pred. No. 7.9;
Matches 33; Conservative 21; Mismatches 59; Indels 18; Gaps 4;

QY 47 KANSEACR-----DGLRAVNECRNVTHLLQELTEAKGQDYDAQATCNHTVMALM 99
DB 1366 KANSEVAQWRKYETDAIOREELAEAKKKLAQRLQEAEEHVAANACASLEKTKORIQ 1425
QY 100 -----ASLDAEKAQG-----OKKVEELGEIETLHMKIODASAEVERLRRENOVSVRIA 149
DB 1426 NEVEDLMDLVERSNACALDKKORNFDPKVSSEMKRYETQAELEASKESSRLSTELF 1485
QY 150 D-KKYPSSQD 159
DB 1486 KVKNYEESLD 1496

RESULT 14
MYHB_HUMAN STANDARD: PRT; 1972 AA.
AC P35749; 000396; P78422; 094944;
DT 01-JUN-1994 (Rel. 29, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Myosin heavy chain, smooth muscle isoform (SMHMC).
GN MYH1 OR KIAA0866.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OC NCBI_TaxID=9606;
RN 11
RP SEQUENCE FROM N.A.
RX MEDLINE=99425270; PubMed=10493829;
RA Lofthus B.J., Kim U.-J., Sneedon V.P., Kalush F., Brandon R.,
RA Fuhmann J., Mason T., Crosby M.L., Barnstead M., Cronin L.,
RA Deslattes Mays A., Cao Y., Xu R.X., Xu R.X., Kang H.-L., Mitchell S.,
RA Eichler E.E., Harris P.C., Venter J.C., Adams M.D.,
RT "Genome duplications and other features in 12 Mb of DNA sequence from
human chromosome 16p and 16q.";
RL Genomics 60:295-308(1999).
[2]
RP SEQUENCE OF 1-1266 FROM N.A.
RC TISSUE-Brain;
RX MEDLINE=99156230; PubMed=10048485;

RA Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Hirosewa M.,
RA Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XII.
RT The complete sequences of 100 new cDNA clones from brain which code
RT for large proteins in vitro.";
RL DNA Res. 5:355-364(1998).
RN [3]
RP SEQUENCE OF 885-1972 FROM N.A.
RX MEDLINE-93263189; PubMed-7684189;
RA Matsuo R., Yoshida M.C., Furutani Y., Imamura S., Kanda N.,
RA Taniguchi M., Masaki T., Takao A.;
RT "Human smooth muscle myosin heavy chain gene mapped to chromosome
RT region 16q12.";
RL Am. J. Med. Genet. 46:61-67(1993).
RN [4]
RP SEQUENCE OF 1093-1972 FROM N.A.
RC TISSUE=Hippocampus;
RA Okajima K.;
RT Submitted (NOV-1992) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: MUSCLE CONTRACTION.
CC -1- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)
CC AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
CC -1- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
CC -1- TISSUE SPECIFICITY: SMOOTH MUSCLE, EXPRESSED IN THE UMBILICAL
CC ARTERY, BLADDER, ESOPHAGUS AND TRACHEA.
CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
CC -1- PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY
CC ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.
CC -1- DISEASE: A CHROMOSOMAL REARRANGEMENT, KNOWN AS PERICENTRIC
CC INVERSION INV(16)(P13Q22), PRODUCES A FUSION PROTEIN THAT CONSISTS
CC OF THE 165 N-TERMINAL RESIDUES OF CBF-BETA (PEP2) WITH THE TAIL
CC REGION OF MYH1. THIS REARRANGEMENT IS ASSOCIATED WITH ACUTE
CC MYELOID LEUKEMIA OF M4EO SUBTYPE.
CC -1- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
CC MEROMYOSIN (LM) AND 1 HEAVY MEROMYOSIN (HM). IT CAN LATER BE
CC SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
CC SUBFRAGMENT (S2).
CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 IQ DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF001548; AAC31665.1; -;
DR EMBL: U91333; AAC35212.1; -;
DR EMBL: AB020673; BAA74889.1; -;
DR EMBL: D10667; -; NOT_ANNOTATED_CDS.
DR EMBL: X69282; CAA49154.1; -;
DR HSSP: P08799; IMN.
DR MIM: 160745; -;
DR InterPro: IPR000048; IQ.
DR InterPro: IPR002928; Myosin_tail.
DR InterPro: IPR002017; Spectrin.
DR InterPro: IPR001609; myosin_head.
DR Pfam: PF00612; IQ; 1.
DR Pfam: PF00063; myosin_head; 1.
DR Pfam: PF01576; myosin_tail; 1.
DR PRINTS: PR00193; MYOSINHEAVY.
DR PRODOM: PD000355; myosin_head; 1.
DR SMART: SM00015; IQ; 2.
DR SMART: SM00242; MYSC; 1.
DR PROSITE: PS50096; IQ; 1.
DR Myosin: Muscle protein; Coiled coil; Thick filament; Actin-binding;
KW Calmodulin-binding; ATP-binding; Methylation; Alkylation;
KW Multigene family; Proto-oncogene; Chromosomal translocation.

FT DOMAIN 1 785 MYOSIN HEAD-LIKE.
FT 786 IQ.
FT DOMAIN 815 COILED COIL (POTENTIAL).
FT 844 1934 CARBOXYL-TERMINAL.
FT 1935 1972 APP (POTENTIAL).
FT NP_BIND 178 185
FT DOMAIN 661 683 ACTIN-BINDING (BY SIMILARITY).
FT 762 776 ACTIN-BINDING (BY SIMILARITY).
FT MOD_RES 129 129 METHYLATION (TR1-) (POTENTIAL).
FT MOD_RES 701 701 ALKYLATION (SH-1) (POTENTIAL).
FT 711 711 ALKYLATION (SH-2) (POTENTIAL).
FT 887 889 EER -> NSE (IN REF. 3).
FT 1263 1266 ELDS -> TLSP (IN REF. 2).
FT 1558 1558 T -> S (IN REF. 3).
FT 1610 1611 KQ -> NE (IN REF. 3).
FT 1786 1786 A -> S (IN REF. 4).
FT 1958 1958 T -> L (IN REF. 3).
SQ SEQUENCE 1972 AA; 227338 MW; 67665B2AEC1277 CRC64;
Query Match 10.7%; Score 95; DB 1; Length 1972;
Best Local Similarity 24.5%; Pred. No. 8;
Matches 27; Conservative 18; Mismatches 55; Indels 10; Gaps 1;
QY 53 CCGCLRAVMECRNVTHLLOELTEAOKGFQDVRAQAATCNHTYALMASLD----- 103
DB 1268 CSDGERARAEALNDKVKHQLQNEVESVTGMLNEAGKAIKLAKDAVSLSSQLODTELLQEE 1327
QY 104 -AKAGOKKVEELEGITFLNKHLDQASAEVRLRENOVLVRIADKK 152
DB 1328 TROKLVSTKRLDLEERNSLDQDLDEMAKONLEHISTLNIQSDSK 1377
RESULT 15
MYHB_RABIT STANDARD; PRT; 1972 AA.
AC P35748;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Myosin heavy chain, smooth muscle isoform (SMHMC).
GN MYH1.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RX SEQUENCE FROM N.A.
RP MEDLINE-92073350; PubMed-1961735;
RA Rabin P., Kelly C., Periasamy M.;
RT "Characterization of a mammalian smooth muscle myosin heavy-chain
RT gene: complete nucleotide and protein coding sequence and analysis of
RT the 5' end of the gene.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:10676-10680(1991).
CC -1- FUNCTION: MUSCLE CONTRACTION.
CC -1- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)
CC AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
CC -1- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
CC -1- PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY
CC ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.
CC -1- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
CC MEROMYOSIN (LM) AND 1 HEAVY MEROMYOSIN (HM). IT CAN LATER BE
CC SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
CC SUBFRAGMENT (S2).
CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 IQ DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2002, 15:25:07 ; Search time 16.33 Seconds
(without alignments)
1059.160 Million cell updates/sec

Title: US-09-828-217-1
Perfect score: 889
Sequence: 1 MASTSYDRCRVPMEDGDKRC.....SSAAPQLLIVLLIGLSALLQ 180

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : PIR-71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	889	100.0	180	2 A56836	bone marrow stroma
2	111	12.5	778	2 T30430	hypothetical prote
3	108	12.1	423	2 T59463	keratin, type I, c
4	107.5	12.1	461	2 H84099	cell wall-binding
5	107	12.0	143	2 G83955	flagellar protein
6	102.5	11.5	1534	2 A56734	ribosome receptor,
7	102	11.5	550	2 A28166	Kupffer cell recep
8	99.5	11.2	304	2 JX0209	lectin, galactose/
9	99	11.1	415	2 S35760	myosin heavy chain
10	98.5	11.1	1963	1 MMKW	myosin heavy chain
11	98	11.0	959	2 A55913	transcytosis-assoc
12	98	11.0	1938	1 MMKW1	myosin heavy chain
13	98	11.0	1938	2 A59293	skeletal myosin be
14	97	10.9	286	2 S07533	puff II/9A-2 prote
15	96.5	10.9	597	2 S40998	hypothetical prote
16	96.5	10.9	893	2 G88551	protein M01A8.2 I
17	96	10.8	286	2 S07532	puff II/9-1 protei
18	95.5	10.7	1937	2 S76705	hypothetical prote
19	95	10.7	1937	2 T38055	myosin heavy chain
20	95	10.7	1972	1 A41604	myosin heavy chain
21	94.5	10.6	1388	2 S74245	serine/threonine-s
22	94	10.6	676	2 S00084	myosin heavy chain
23	93.5	10.5	140	2 H64629	hypothetical prote
24	93.5	10.5	1938	1 S06005	myosin alpha heavy
25	93.5	10.5	1939	2 T48175	myosin heavy chain
26	93	10.5	359	2 T50712	TOP AP - chicken
27	93	10.5	848	2 A44972	paramyosin - nemat
28	93	10.5	879	2 A48575	paramyosin - nemat
29	93	10.5	1938	2 J05421	smooth muscle myos

30	93	10.5	1972	2 J05420	smooth muscle myos
31	92.5	10.4	392	2 G95258	secreted 45 kd pro
32	92.5	10.4	392	2 B88124	general stress pro
33	92.5	10.4	1509	1 A27224	myosin heavy chain
34	92.5	10.4	1938	2 T49464	alpha cardiac myos
35	92	10.3	746	2 T47237	myosin II heavy ch
36	92	10.3	866	2 S04027	paramyosin - Caeno
37	92	10.3	872	2 T19296	hypothetical prote
38	91.5	10.3	244	2 T16913	beta-myosin heavy
39	91.5	10.3	764	2 I51302	myosin heavy chain
40	91.5	10.3	1039	2 S18199	myosin heavy chain
41	91.5	10.3	1934	2 T48153	myosin heavy chain
42	91.5	10.3	1935	1 A37102	myosin beta heavy
43	91.5	10.3	1935	1 S06006	myosin beta heavy
44	91.5	10.3	1939	1 A46762	myosin alpha heavy
45	91.5	10.3	2442	2 T08621	centrosome associa

ALIGNMENTS

```
RESULT 1
A56836
bone marrow stromal cell surface protein BST-2 - human
C:Species: Homo sapiens (man)
C:Date: 11-Aug-1995 #sequence_revision 11-Aug-1995 #text_change 21-Jul-2000
C:Accession: A56836
R: Ishikawa, J.; Kaisho, T.; Tomizawa, H.; Lee, B.O.; Kobune, Y.; Inazawa, J.; Oritani
A:Title: Molecular cloning and chromosomal mapping of a bone marrow stromal cell surf
A:Reference number: A56836; MUID:95331788
A:Accession: A56836
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-180 <ISH>
A:Cross-references: GB:D8137; NID:9457563; PIDN:BA05679.1; PID:9506861
C:Genetics:
A:Gene: GDB:BSR2
A:Cross-references: GDB:409946; OMIM:600534
A:Map position: 19p13.2-19p13.2
C:Keywords: transmembrane protein

Query Match      100.0%; Score 889; DB 2; Length 180;
Best local Similarity 100.0%; Pred. No. 2.6e-62;
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASTSYDRCRVPMEDGDKRCKLLIGILVLLIIVITGVPLITFIKANSEACRDLRAV 60
    |||
DB 1 MASTSYDRCRVPMEDGDKRCKLLIGILVLLIIVITGVPLITFIKANSEACRDLRAV 60
    |||

QY 61 MECRNVTLLLOELTEAOKGFQDVEAQAATCNHTVMALMASIDAKAOCQKVELEGEI 120
    |||
DB 61 MECRNVTLLLOELTEAOKGFQDVEAQAATCNHTVMALMASIDAKAOCQKVELEGEI 120
    |||

QY 121 TTLNKHLDASAEVERLRENOVLSVRADKKYRSSQDSSSAAPQLLIVLLIGLSALLQ 180
    |||
DB 121 TTLNKHLDASAEVERLRENOVLSVRADKKYRSSQDSSSAAPQLLIVLLIGLSALLQ 180
    |||

RESULT 2
T30430
hypothetical protein ORF82 - Lymantria dispar nuclear polyhedrosis virus
C:Species: Lymantria dispar nuclear polyhedrosis virus, LdMNPV
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C:Accession: T30430
R: Kuzio, J.; Pearson, M.N.; Harwood, S.H.; Funk, C.J.; Evans, J.T.; Slavicek, J.M.; R
A:Title: Sequence and analysis of the genome of a baculovirus pathogenic for Lymantri
A:Reference number: Z20836; MUID:99124785
A:Accession: T30430
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
```

A:Residues: 1-778 <KU2>
A:Cross-references: EMBL:AF081810; PIDN:AACT0268.1

Query Match 12.5%; Score 111; DB 2; Length 778;
Best Local Similarity 32.5%; Pred. No. 0.48;
Matches 38; Conservative 22; Mismatches 47; Indels 10; Gaps 4;

QY 46 IKANSEACRDLRAVMECR-NVTNLLQDELTEAOKGFQDVEAQAATCNHTVMAALMSLDA 104
DB 520 LKAKSELND-LQAKAEADQANARLQAEIDSKR-----AESADADLRNVLAQLEAEAS 574

QY 105 EKAQGGKVEELEGITTLNHLKQDASAEVERLRRENOVLVRADKKYPPSSODSS 161
DB 575 GAADLQNRARLRLEASSGLTRRQESAAEVLLALRROKEDLERRTAAS----AAQDVS 627

RESULT 3
59463
A:Title: keratin, type I, cytoskeletal - mouse
A:Alternate names: endo B cytokeatrin; keratin D
C:Species: Mus musculus (house mouse)
C:Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 10-Dec-1999
C:Accession: I59463; A25621; A28428; J10406
R:Alonso, A.; Weber, T.; Jorcano, J.L.
Roux's Arch. Dev. Biol. 196, 16-21, 1987
A:Title: Cloning and characterization of keratin D, a murine endodermal cytoskeletal pro
A:Reference number: I59463
A:Accession: I59463
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-423 <RES>
A:Cross-references: GB:M36376; NID:q198587; PIDN:AAA39373.1; PID:q293682
R:Singer, P.A.; Trevor, K.; Oshima, R.G.
J. Biol. Chem. 261, 538-547, 1986
A:Title: Molecular cloning and characterization of the endo B cytokeatrin expressed in F
A:Reference number: A25621; M0ID:86085876
A:Accession: A25621
A:Molecule type: mRNA
A:Residues: 1-243; 'D', 245-252; 'A', 254-423 <SIN>
A:Cross-references: GB:M1686; NID:q198620; PIDN:AAA39390.1; PID:q293685
R:Oshima, R.G.; Trevor, K.; Shevinsky, L.H.; Ryder, O.A.; Cecena, G.
Genes Dev. 2, 505-516, 1988
A:Title: Identification of the gene coding for the endo B murine cytokeatrin and its me
A:Reference number: A28428; M0ID:88255838
A:Accession: A28428
A:Molecule type: DNA
C:Residues: 1-113 <OSH>
A:Cross-references: GB:Y00217; NID:q50842; PIDN:CAA68365.1; PID:q50843
R:Chino, Y.; Morita, T.; Zhang, F.; Srimahasongram, S.; Tondella, M.L.C.; Matsumoto,
Gene 70, 85-95, 1988
A:Title: Nucleotide sequence and structure of the mouse cytokeatrin endoB gene.
A:Reference number: J10406; M0ID:89196920
A:Accession: J10406
A:Molecule type: DNA
A:Residues: 1-113; 'F', 135-243; 'D', 245-252; 'A', 254-423 <ICH>
A:Cross-references: GB:M2832; NID:q340757; PIDN:AAA37552.1; PID:q532610
C:Genetics:
A:Gene: endoB; KERD
A:Introns: 132/3; 160/2; 212/3; 267/3; 309/3; 384/2
C:superfamily: cytoskeletal keratin
C:keywords: coiled coil; intermediate filament

Query Match 12.1%; Score 108; DB 2; Length 423;
Best Local Similarity 28.0%; Pred. No. 0.44;
Matches 52; Conservative 28; Mismatches 56; Indels 50; Gaps 8;

QY 44 FTIKANSE-ACRD-----GLRAVMECRNVTHL-----LQDELTEAOKG----FQD 83
DB 157 FRAVYETELARQSVESDINGLRKRVVDJNTNRLQLETEALKEKELLFMKKNHEEVQ 216

QY 84 VEAQAATCNHTV-----MALMASLDAE-KAOGOKVEELE-----GEIT 121

DB 217 LKQIASSGTLVEVDAPKSDLSKIMANIRAOYEALGQKREELDKYWSQJIEESTVYT 276

QY 122 TLNHRKQDASAEVERLRRENOVLVRADKKYPPSSODSS-----SAAPOLLIVL 173
DB 277 TSSAEIRDETTTLTELRKLTQLEIDLSKMNONTLENSLDGVEARKYAKHQLNGVLL 336

QY 174 GLSALL 179
DB 337 HLESEL 342

RESULT 4
H84099
A:Title: cell wall-binding protein BH3600 [imported] - Bacillus halodurans (strain C-125)
C:Species: Bacillus halodurans
C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C:Accession: H84099
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; H
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans a
A:Reference number: A83650; M0ID:20512582; PMID:11058132
A:Accession: H84099
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-461 <STO>
A:Cross-references: GB:AP001519; GB:BA000004; NID:g10176109; PIDN:BAB07319.1; GSPDB:G
A:Experimental source: strain C-125
C:Genetics:
A:Gene: BH3600

Query Match 12.1%; Score 107.5; DB 2; Length 461;
Best Local Similarity 26.1%; Pred. No. 0.52;
Matches 31; Conservative 28; Mismatches 55; Indels 5; Gaps 2;

QY 33 IYVILGVPLIIFIRKANSEACRDLRAVMECRNVTHLQDELTEAOKGFQDVEAQAATCN 92
DB 5 ISLVAAAGLTFSLFSQSIEDA-KANSLQQISDVOKEROKOKEAEV----- 59

QY 93 HTVMAALMSLDAEKAOGOKVEELEGITTLNHLKQDASAEVERLRRENOVLVRADK 151
DB 60 KEVERKELGDTAIEIRLDEKVEETSKIQEKREIEVQAEIIEELKEIIEERIAER 118

RESULT 5
G83955
A:Title: flagellar protein required for flagellar formation flil [imported] - Bacillus halodur
C:Species: Bacillus halodurans
C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C:Accession: G83955
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; H
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans a
A:Reference number: A83650; M0ID:20512582; PMID:11058132
A:Accession: G83955
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-143 <STO>
A:Cross-references: GB:AP001515; GB:BA000004; NID:g10174886; PIDN:BAB06166.1; GSPDB:G
A:Experimental source: strain C-125
C:Genetics:
A:Gene: flil

Query Match 12.0%; Score 107; DB 2; Length 143;
Best Local Similarity 22.1%; Pred. No. 0.17;
Matches 31; Conservative 34; Mismatches 49; Indels 26; Gaps 4;

QY 23 LIGIGILVLLIIVLGVPLIIFIRKANSEACRDLRAVMEC-----RNVTHLQDEL 74
DB 6 LVNIMLIIIVLTVGVAVLIFVNFVNEDEQREPTIDEIRIAQSYETERTITNLSNDF 65

J. Bacteriol. 174, 4967-4976, 1992
A:Title: Architecture of the vir regulons of group A streptococci parallels opacity fact
A:Reference number: A42711; MUID:92332431
A:Accession: A42711
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 343-415 <HAA>
A:Cross-references: GB:M86806; NID:g153630; PIDN:AAA26887.1; PID:g153631
A:Experimental source: strain CS101, OF+
A:Note: sequence extracted from NCBI backbone (NCBIN:108942, NCBI:P:108945)
C:Superfamily: M5 protein

Query Match 11.1%; Score 99; DB 2; Length 415;
Best Local Similarity 26.9%; Pred. No. 2.2;
Matches 35; Conservative 29; Mismatches 46; Indels 20; Gaps 5;

QY 57 LRVMQERN-VTHLQOELTEAOKGQDVEAQAAT-----CNHYMALMSLD---A 104
184 VKSQLEKKNKEIDELKQODASKTEE-TANLQSEATLENLGSAKHELTDLQAKLDATA 242
QY 105 EKAQOGKVEELEG-----ETTLNHLKLDASAEVERLRRENOVLSVRIDAKKYPSS 157
DB 243 EKAKLESQETLENLGSAKRELTDLQAKLDANAEKELQSQAAALEKQLAFTKRELAD 302
QY 158 QDSSSAAPQ 167
DB 303 LQAKLATNQ 312

RESULT 10
MMKW
myosin heavy chain B [similarity] - Caenorhabditis elegans
N:Contains: myosin ATPase (EC 3.6.1.32)
C:Species: Caenorhabditis elegans
C:Date: 13-Jun-1983 #sequence revision 19-May-2000 #text change 19-Jan-2001
C:Accession: T20770; T21629; A93958; A93287; A21074; A02992
R:Kershaw, J.
submitted to the EMBL Data Library, November 1996
A:Reference number: Z19322
A:Accession: T20770
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1963 <W12>
A:Cross-references: EMBL:Z8107; PIDN:CAB05505.1; GSPDB:GN00019; CESP:F11C3.3
A:Experimental source: clone F32A7
R:Karn, J.; Brenner, S.; Barnett, L.
Proc. Natl. Acad. Sci. U.S.A. 80, 4253-4257, 1983
A:Title: Protein structural domains in the Caenorhabditis elegans unc-54 myosin heavy ch
A:Reference number: A93958; MUID:83273600
A:Accession: A93958
A:Molecule type: DNA
A:Residues: 1-61, 'EMSVIR', 65-376, 'V', 378-1963 <KAR>
A:Cross-references: GB:J01050; NID:g156399; PIDN:AAA8124.1; PID:g156400
R:McLachlan, A.D.; Karn, J.
Nature 299, 226-231, 1982
A:Title: Periodic charge distributions in the myosin rod amino acid sequence match cross
A:Reference number: A93287; MUID:82372395
A:Accession: A93287
A:Molecule type: DNA
A:Residues: 847-1333, 'R', 1335-1876, 'L', 1878-1963 <MC1>
R:Wills, N.; Gesteland, R.F.; Karn, J.; Barnett, L.; Bolten, S.; Waterston, R.H.
Cell 33, 575-583, 1983
A:Title: The genes sup-7 X and sup-5 III of Caenorhabditis elegans suppress amber nonser
A:Reference number: A21074; MUID:83232892
A:Accession: A21074
A:Molecule type: DNA

A:Residues: 1873-1963 <W13>
A:Cross-references: GB:V01494; GB:J01049; NID:g6783; PIDN:CAA24738.1; PID:g6784
C:Genetics:
A:Gene: unc-54; CESP:F11C3.3
A:Map position: 1
A:Introns: 21/3; 64/3; 111/3; 264/1; 525/3; 951/2; 1747/3; 1819/3; 1894/3
C:Superfamily: myosin heavy chain; myosin motor domain homology
C:Keywords: actin binding; ATP; coiled coil; hydrolase; methylated amino acid; muscle
F:84-775/Domain: myosin motor domain homology <MMOT>
F:174-181/Region: nucleotide-binding motif A (P-loop)
F:662-684/Region: actin binding #status predicted
F:766-780/Region: actin binding #status predicted
F:848-1963/Domain: coiled coil #status predicted <CO1>
F:848-1162/Region: 52
F:1163-1963/Region: light meromyosin
F:125/Modified site: N6,N6,N6-trimethyllysine (Lys) #status predicted
F:180/Binding site: ATP (Lys) #status predicted
F:702,712/Active site: Cys #status predicted

Query Match 11.1%; Score 98.5; DB 1; Length 1963;
Best Local Similarity 24.7%; Pred. No. 12;
Matches 36; Conservative 21; Mismatches 58; Indels 31; Gaps 5;

QY 47 KANSEACRQGLRAVMCRNVTLLQOELTEAOKGQDVEAQAATCNHYMALMSIDAER 106
DB 1808 EAEAAALGCGKVIARLEORVRELESELDEGRFODANKLGRADRRVRELQFQVDEPK 1867
QY 107 AQ-----GQKVEELEGETTLN-----HKLDASAEVERLRREN 141
DB 1868 KNFERLQDLIDKQKIKTKQKQVEER-ELANLNLYKTKLTHLEDAEERAD--QAEV 1924
QY 142 QVLSVRI---ADKKYPPSSODSSSA 164
DB 1925 SLSKMSKSRASVAPRGQSSKSA 1950

RESULT 11
A55913
transcytosis-associated protein p115 - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 23-Mar-1995 #sequence revision 05-Apr-1995 #text change 05-Nov-1999
C:Accession: A55913
R:Barroso, M.; Nelson, D.S.; Szul, E.
Proc. Natl. Acad. Sci. U.S.A. 92, 527-531, 1995
A:Title: Transcytosis-associated protein (TAP)/p115 is a general fusion factor requir
A:Reference number: A55913; MUID:95132633
A:Accession: A55913
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-959 <BAR>
A:Cross-references: GB:U15589; NID:g558474; PIDN:AAC52151.1; PID:g558475
C:Keywords: membrane fusion; membrane trafficking

Query Match 11.0%; Score 98; DB 2; Length 959;
Best Local Similarity 30.1%; Pred. No. 6.1;
Matches 37; Conservative 21; Mismatches 57; Indels 8; Gaps 3;

QY 60 VMCRNVTLLQOELTEAOKGQDVEAQAATCNHYMALMSIDAERKQKVEELEG 118
DB 739 IEELRSHQVLLQSLAEKPTVLENLSSQVSGSEQALTCSPRAE-----QVAELKQ 792
QY 119 EITTLNHLKLDASAEVERLRRENOVLSVRIDAKKYPSSODSSSAAPQDLITLGLSA 177
DB 793 ELKALKSQLSQSLSEITRLOTENRELQGRARFLAKSVPEGSELYTAKTIDYERLSA 852
QY 178 LQD 180
DB 853 LQD 855

RESULT 12

MMKWI
 myosin heavy chain D [similarity] - Caenorhabditis elegans
 N:Alternate names: myosin heavy chain I
 N:Contains: myosin ATPase (EC 3.6.1.32)
 C:Species: Caenorhabditis elegans
 C:Date: 28-Feb-1986 #sequence_revision 19-May-2000 #text_change 19-Jan-2001
 C:Accession: T21193; T23973; S02772; A02993
 R:McMurray, A.
 submitted to the EMBL Data Library, April 1996
 A:Reference number: Z19388
 A:Accession: T21193
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1938 <M12>
 A:Cross-references: EMBL:Z71261; PIDN:CAA95806.1; GSPDB:GN00019; CESP:R06C7.10
 A:Experimental source: clone F21C3
 R:Gardner, A.
 submitted to the EMBL Data Library, April 1996
 A:Reference number: Z19825
 A:Accession: T23973
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1938 <M12>
 A:Cross-references: EMBL:Z71266; PIDN:CAA95848.1; GSPDB:GN00019; CESP:R06C7.10
 A:Experimental source: clone R06C7
 R:Diib, N.J.; Maruyama, I.N.; Krause, M.; Karn, J.
 J. Mol. Biol. 205, 603-613, 1989
 A:Title: Sequence analysis of the complete Caenorhabditis elegans myosin heavy chain gen
 A:Reference number: S02771; MUID:89178677
 A:Accession: S02772
 A:Status: nucleic acid sequence not shown
 A:Molecule type: DNA
 A:Residues: 1-376, 'V', 378-390, 'V', 392-576, 'L', 578-680, 'R', 682-1938 <DIB>
 A:Cross-references: EMBL:X08065; NID:96785; PIDN:CAA30854.1; PID:96786
 R:Karn, J.; Brenner, S.; Barnett, L.
 Proc. Natl. Acad. Sci. U.S.A. 80, 4253-4257, 1983
 A:Title: Protein structural domains in the Caenorhabditis elegans unc-54 myosin heavy ch
 A:Reference number: A93958; MUID:83273600
 A:Accession: A02993
 A:Molecule type: DNA
 A:Residues: 24-93, 'E', 95-97, 'R', 99-376, 'V', 378-388, 'GDV', 392-407, 'N', 409-473, 'G', 475-576
 C:Genetics:
 A:Gene: myo-1; CESP:R06C7.10
 A:Map position: 1
 A:Introns: 23/3; 114/3; 229/1; 264/1; 320/1; 857/3; 1745/3; 1814/1; 1892/3
 C:Superfamily: myosin heavy chain; myosin motor domain homology
 C:Keywords: actin binding; ATP; coiled coil; hydrolase; methylated amino acid; muscle cc
 F:87-773/Domain: myosin motor domain homology <MMOT>
 F:177-184/Region: nucleotide-binding motif A (P-loop)
 F:660-682/Region: actin binding #status predicted
 F:764-778/Region: actin binding #status predicted
 F:846-1938/Domain: coiled coil #status predicted <COI>
 F:846-1160/Region: S2
 F:1161-1938/Region: light meromyosin
 F:128/Modified site: N6,N6-trimethyllysine (Lys) #status predicted
 F:183/Binding site: ATP (Lys) #status predicted
 F:700/710/Active site: Cys #status predicted

Query Match 11.0%; Score 98; DB 1; Length 1938;
 Best Local Similarity 26.6%; Pred. No. 13;
 Matches 34; Conservative 24; Mismatches 58; Indels 12; Gaps 4;

QY 51 EACRQGLRAVME-----CRVTHLQO--ELTEAQ-KGFQDVEQAATCNHTVAMLASID 103
 DB 1324 KAAEELHERQEPHACKNLEHLLQCHLEEQINGKDDIQRLSRINSEISQWARYE 1383
 QY 104 AAKAOGKQVELEGEITLNNKLQDASAEVERLRENOVLVSRIDKKYPPSSODSSA 163
 DB 1384 GEGVGSSELEELKKQKRMVNDLQALSA-----QNVISLEKAKGLIAETEDARSD 1438
 QY 164 AAPOLLIV 171

DB 1439 VDRHLTVI 1446

RESULT 13
 A59293
 skeletal myosin heavy chain - domestic rabbit
 C:Species: Oryctolagus cuniculus
 C:Date: 09-Jun-2000 #sequence_revision 09-Jun-2000 #text_change 08-Sep-2000
 C:Accession: A59293
 R:Maeda, K.; Hostilova, E.; Roess, Kleinkauf, A.; Schuster, H.; Gasperik, J.; Wittlingh
 submitted to GenBank, July 1995
 A:Description: Isolation, sequencing of myosin heavy chain cDNA from rabbit skeletal
 A:Reference number: A59293
 A:Accession: A59293
 A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 1-1938 <MAE>
 A:Cross-references: GB:U32574; NID:9940232; PIDN:AAA74199.1; PID:9940233
 A:Experimental source: strain New Zealand White; cell type skeletal muscle fiber type
 C:Genetics:
 A:Gene: MHC
 C:Superfamily: myosin heavy chain; myosin motor domain homology
 F:89-769/Domain: myosin motor domain homology <MMO>

Query Match 11.0%; Score 98; DB 2; Length 1938;
 Best Local Similarity 26.1%; Pred. No. 13;
 Matches 37; Conservative 20; Mismatches 45; Indels 40; Gaps 5;

QY 47 KANSEACR-----DGLRAVMECRNVTHLQOELTEAQKGFQDVEQAATCNHTVAM 99
 DB 1366 KANSEVAQWRTRYEDIDAIRTELEAKKKIAQRLQDAE---EHVEAVAKC----- 1414
 QY 100 ASLDAEAKQKQVELEGEITTLN-----HKQDASAEVERLR 138
 DB 1415 ASLEKQKQRLQNEVEDLMDVERTNAACAALDKQRNPDKILAEWKHKYEETHAELEAQ 1474
 QY 139 REMOVLVRIAD-KKYPPSSOD 159
 DB 1475 KESRSLSTEVFKVKNAYESLD 1496

RESULT 14
 S07533
 puff II/9A-2 protein precursor - fungus gnat (Sciara coprophila)
 C:Species: Sciara coprophila
 C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 20-Jun-2000
 C:Accession: S07533
 R:DiBartolomeis, S.M.; Gerbl, S.A.
 J. Mol. Biol. 210, 531-540, 1989
 A:Title: Molecular characterization of DNA puff II/9A genes in Sciara coprophila.
 A:Reference number: S07532; MUID:90133907
 A:Accession: S07533
 A:Status: not compared with conceptual translation
 A:Molecule type: DNA
 A:Residues: 1-286 <DIB>
 A:Cross-references: GB:X51679; NID:910113; PID:q1405812
 C:Genetics:
 A:Map position: II/9A
 C:Keywords: coiled coil; glycoprotein
 F:1-20/Domain: signal sequence #status predicted <SIG>
 F:21-286/Product: puff II/9A protein #status predicted <MAP>
 F:156/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 10.9%; Score 97; DB 2; Length 286;
 Best Local Similarity 25.6%; Pred. No. 2.1;
 Matches 31; Conservative 24; Mismatches 36; Indels 30; Gaps 5;

QY 46 IKANSEACRQGLRAVMECR-RNVTHL-----LQOELTEAQKGFQDVEQAATCNHTVAM 98
 DB 89 LKREKARQKAEKAEKQKNTENIKETIEQLKELEAEQAKLEKCKKEELADCK----- 142

OY 99 MASIDAERAGOCRKVEELEGITTLNHNKLO-----DASAEYERLRE-----NOVLAV 146
/ | : ||| | | | : | : : : : :
Db 143 -----KENAKLLNNIEBELNCTITQLOEKLEBRCRGRENDIQCQLDECKKKKNICNNNELINC 197

OY 147 R 147
Db 198 R 198

RESULT 15

hypothetical protein M01A8.2 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 06-Jan-1995
C:Accession: S40998
R:Hawkins, T.; Thomas, K.
submitted to the EMBL Data Library, October 1993
A:Reference number: S40997
A:Accession: S40998
Status: preliminary
Molecule type: DNA
A:Residues: 1-597 <HAN>
A:Cross-references: EMBL:227081
C:Genetics:
A:Introns: 40/3; 76/1; 125/1; 199/1; 240/1; 274/1; 292/3; 392/3; 429/2; 486/2; 521/1

	Query Match	10.9%	Score 96.5	DB 2	length 597	
	Best Local Similarity	24.8%	Pred. No. 4.9			
	Matches 32	Conservative 29	Mismatches 39	Indels 29	Gaps 4	
QY	48 ANSEACRDLGRAVMECRNVTHLQ--QELTEAKGFQDY----EQAQATCNHTVALMAS 101					
Db	270 SNOGYIRHMANV-ESLQKTHETQIAEKKNKEEFERNEEERARREAEVCAAMNRHQKVAC 328					
QY	102 LPAEAQAGQKKKEELE-----GTTLNHKLQDASAEYERLR 139					
Db	329 LDEKSEAEKQCEQLVNDKKVLAALANDCDHRNOMLTKEISSLOTALEMKSAEMKEKRO 388					
QY	140 ENQVLSVRI 148					
Db	389 KNQNTSLQV 397					

Search completed: June 4, 2002, 15:27:21
Job time: 134 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OW protein - protein search, using sw model

Run on: June 4, 2002, 15:24:47 ; Search time 13.06 Seconds
(without alignments)
336.647 Million cell updates/sec

Title: US-09-828-217-1

Perfect score: 889
Sequence: 1 MASTSYDYCRVPMEDGDKRC.....SSAAPQLLIVLLGLSALIQ 180

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PTCUS_COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	889	100.0	180	2 US-08-624-650-1	Sequence 1, Appli
2	99.5	11.2	270	2 US-09-055-095-4	Sequence 4, Appli
3	99.5	11.2	270	2 US-08-809-494-2	Sequence 2, Appli
4	99.5	11.2	270	4 US-09-352-302-2	Sequence 2, Appli
5	99.5	11.2	273	2 US-08-809-494A-4	Sequence 4, Appli
6	99.5	11.2	273	4 US-09-352-302-4	Sequence 4, Appli
7	95	10.7	885	2 US-08-533-306A-4	Sequence 4, Appli
8	95	10.7	885	2 US-08-742-923A-4	Sequence 4, Appli
9	93.5	10.5	1886	4 US-08-938-105-3	Sequence 4, Appli
10	92.5	10.4	292	2 US-08-688-342-4	Sequence 4, Appli
11	92.5	10.4	292	2 US-09-113-788-4	Sequence 4, Appli
12	89.5	10.1	316	4 US-09-111-470-4	Sequence 4, Appli
13	89.5	10.1	1939	4 US-09-310-167A-1	Sequence 1, Appli
14	88	9.9	288	3 US-08-312-949-4	Sequence 4, Appli
15	88	9.9	288	3 US-08-446-201-4	Sequence 4, Appli
16	88	9.9	619	1 US-08-465-746-2	Sequence 2, Appli
17	88	9.9	619	1 US-08-214-164-2	Sequence 2, Appli
18	88	9.9	619	2 US-08-467-852A-3	Sequence 3, Appli
19	88	9.9	619	2 US-08-246-636-2	Sequence 3, Appli
20	88	9.9	619	2 US-08-247-491A-3	Sequence 3, Appli
21	88	9.9	619	2 US-08-319-795-2	Sequence 2, Appli
22	88	9.9	619	2 US-08-468-985-2	Sequence 2, Appli
23	88	9.9	619	3 US-08-312-949-2	Sequence 2, Appli
24	88	9.9	648	1 US-08-072-070-2	Sequence 2, Appli
25	88	9.9	648	1 US-08-469-434-2	Sequence 2, Appli
26	88	9.9	648	1 US-08-214-222-2	Sequence 2, Appli
27	88	9.9	648	2 US-08-467-852A-2	Sequence 2, Appli

28	88	9.9	648	2 US-08-468-718-2	Sequence 2, Appli
29	88	9.9	648	2 US-08-247-491A-2	Sequence 2, Appli
30	88	9.9	648	3 US-08-446-201-3	Sequence 2, Appli
31	88	9.9	695	1 US-08-127-499A-23	Sequence 23, Appli
32	88	9.9	695	1 US-08-482-847-23	Sequence 23, Appli
33	87.5	9.8	289	1 US-08-072-070-4	Sequence 4, Appli
34	87.5	9.8	289	1 US-08-469-434-4	Sequence 4, Appli
35	87.5	9.8	289	1 US-08-214-222-4	Sequence 4, Appli
36	87.5	9.8	289	2 US-08-467-852A-5	Sequence 5, Appli
37	87.5	9.8	289	2 US-08-468-718-4	Sequence 4, Appli
38	87.5	9.8	289	2 US-08-247-491A-5	Sequence 5, Appli
39	87.5	9.8	477	1 US-08-402-217A-3	Sequence 3, Appli
40	87.5	9.8	477	1 US-08-700-178-3	Sequence 3, Appli
41	87.5	9.8	477	3 US-08-995-654-3	Sequence 3, Appli
42	87.5	9.8	667	4 US-09-071-709-9	Sequence 9, Appli
43	85.5	9.6	459	4 US-09-071-709-1	Sequence 1, Appli
44	85	9.6	344	6 5210183-2	Patent No. 5210183
45	85	9.6	683	6 5210183-3	Patent No. 5210183

ALIGNMENTS

RESULT 1
US-08-624-650-1
Sequence 1, Application US/08624650
Patent No. 5914252
GENERAL INFORMATION:
APPLICANT: HIRANO, TOSHIO
APPLICANT: KAISHO, TSUNEYASU
TITLE OF INVENTION: MEMBRANE PROTEIN POLYPEPTIDE HAVING
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: P. C.
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/624,650
FILING DATE: 22-MAY-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/JP94/01732
FILING DATE: 14-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 5-281622
FILING DATE: 15-OCT-1993
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 7625-001-0 PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-3220
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 180 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-624-650-1

Query Match 100.0%; Score 889; DB 2; Length 180;
Best Local Similarity 100.0%; Pred. No. 4.3e-88;
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASTSDYCVPMEDGKCKLLGILVLLIIVLGVPLLIITRKANSEACRDGLRAV 60
DB 1 MASTSDYCVPMEDGKCKLLGILVLLIIVLGVPLLIITRKANSEACRDGLRAV 60
QY 61 MECRNVTHTLQOELTEAOKGQFQVEAQAATCNHTVMAALMSLDAEKAGQKVEELEG 120
DB 61 MECRNVTHTLQOELTEAOKGQFQVEAQAATCNHTVMAALMSLDAEKAGQKVEELEG 120
QY 121 TITLHKLQDASAEVERLRNRQVLSVRIADKRTYPPSSQDSSAAAPOLLIVLGLSALLQ 180
DB 121 TITLHKLQDASAEVERLRNRQVLSVRIADKRTYPPSSQDSSAAAPOLLIVLGLSALLQ 180

RESULT 2
US-09-055-095-4
Sequence 4, Application US/09055095
Patent No. 5945308
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Paterson, Chandra
APPLICANT: Cotley, Neil C.
TITLE OF INVENTION: HUMAN OXIDIZED LDL RECEPTOR
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Dr.
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/055,095
FILING DATE: Filed Herewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0500 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 270 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 1902982
US-09-055-095-4

Query Match 11.2%; Score 99.5; DB 2; Length 270;
Best Local Similarity 25.9%; Pred. No. 0.0068;
Matches 36; Conservative 25; Mismatches 41; Indels 37; Gaps 5;

QY 23 LIGIGILVLLIIVLGVPLLIITRKANSEACRDGLRAVMECRNVTHTLQOELTEAOKGQ 82
DB 40 VLICGLIVTVILLLILOVSDLIKQO-----ANITH--QEDILEGO----- 80

QY 83 DVEAQAATCNHTVMAALMSLDAEKAGQKVEELEGITTLNHLKQDASAEVERLRNRQ 142
DB 81 -----ILAQRSEKS--AQESQKELKEIETLAKHLDEKSKKLMELHQNQL 124
QY 143 VLSVRIADKRTY--PSSQD 159
DB 125 NLOEVLKEAANYSGPCPD 143

RESULT 3
US-08-809-494A-2
Sequence 2, Application US/08809494A
Patent No. 5962260
GENERAL INFORMATION:
APPLICANT: Masaki, Tatsuya
APPLICANT: Sawamura, Tomoo
TITLE OF INVENTION: Modified Low-Density Lipoprotein
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: McAlay Fisher Nissen Goldberg & Kiel
STREET: 261 Madison Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10016-2391
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/809,494A
FILING DATE: 24-MAR-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-321705
FILING DATE: 30-NOV-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 7-214206
FILING DATE: 31-JUL-1995
ATTORNEY/AGENT INFORMATION:
NAME: Goldberg, Jules E
REGISTRATION NUMBER: 24408
REFERENCE/DOCKET NUMBER: JG-YV-4363PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 986-4090
TELEFAX: 212 818-9479
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 270 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-809-494A-2

Query Match 11.2%; Score 99.5; DB 2; Length 270;
Best Local Similarity 25.9%; Pred. No. 0.0068;
Matches 36; Conservative 25; Mismatches 41; Indels 37; Gaps 5;

QY 23 LIGIGILVLLIIVLGVPLLIITRKANSEACRDGLRAVMECRNVTHTLQOELTEAOKGQ 82
DB 40 VLICGLIVTVILLLILOVSDLIKQO-----ANITH--QEDILEGO----- 80
QY 83 DVEAQAATCNHTVMAALMSLDAEKAGQKVEELEGITTLNHLKQDASAEVERLRNRQ 142
DB 81 -----ILAQRSEKS--AQESQKELKEIETLAKHLDEKSKKLMELHQNQL 124
QY 143 VLSVRIADKRTY--PSSQD 159
DB 125 NLOEVLKEAANYSGPCPD 143

```

RESULT      4
US-09-352-302-2
; Sequence 2, Application US/09352302
; Patient No. 6197937
; GENERAL INFORMATION:
; APPLICANT: Sawamura, Tatsuya
; APPLICANT: Masaki, Tomoo
; TITLE OF INVENTION: Modified Low-Density Lipoprotein
; TITLE OF INVENTION: Receptor
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McAnulty Fisher Nissen Goldberg & Kiel
; STREET: 261 Madison Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10016-2391
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/352,302
; FILING DATE: 12-JUL-1999
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 6-321705
; FILING DATE: 30-NOV-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 7-214206
; FILING DATE: 31-JUL-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Goldberg, Jules E
; REGISTRATION NUMBER: 24408
; REFERENCE/DOCKET NUMBER: JG-yy-4363PCT/D
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 986-4090
; TELEFAX: 212 818-9479
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 270 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-352-302-2

Query Match          11.2%; Score 99.5; DB 4; Length 270;
Best local similarity 25.9%; Pred. NO.0.0068;
Matches 36; Conservative 25; Mismatches 41; Indels 37; Gaps 5

QY    23 LGGIGILLIIVILGVPLIIFTIKANSEACRDGRAYWECRNVTLLLOQLLEAQGFQ 82
       :|::||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db    40 VLGGLLVTVLIIQLSGVSLDKKQ-----ANITHT--QEDILEGQ---- 80

QY    83 DVEAQAICNRTVMALMASLSDEAKRQGCKKVVELEGSETTTLNHKKLDADSAEVERLRRNQ 142
       :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db    81 -----LFLQRNRSEKS-AQSOKELKEMETTLAHKIDSKRKIMELHRLNL 124

QY    143 VLSTRIADKKYY--PSSOD 159
       |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db    125 NLGEVLKEANYSGPCPPD 143

RESULT      5
US-08-809-494A-4
; Sequence 4, Application US/08809494A
; Patient No. 5962260
; GENERAL INFORMATION:
; APPLICANT: Sawamura, Tatsuya

```

```

APPLICANT: Masaki, Tomoo
TITLE OF INVENTION: Modified Low-Density Lipoprotein Receptor
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: McAulay Fisher Nissen Goldberg & Kiel
STREET: 261 Madison Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10016-2391
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patentln Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/809,494A
FILING DATE: 24-MAR-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-321705
FILING DATE: 30-NOV-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 7-214206
FILING DATE: 31-JUL-1995
ATTORNEY/AGENT INFORMATION:
NAME: Goldberg, Jules E
REGISTRATION NUMBER: 24408
REFERENCE/DOCKET NUMBER: JG-YY-4363PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 866-4090
TELEFAX: 212 818-9479
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 273 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-809-494A-4

Query Match      11.2%; Score 99.5; DB 2; Length 273;
Best Local Similarity 25.9%; Pred. No. 0.0069;
Matches 36; Conservative 25; Mismatches 41; Indels 37; Gaps 5

QY 23 LIGISILVLIVLIGVPLITFTIKANSEACRDGRVAWECHNNVTLLHQELTEAKGFQ 82
   ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 43 VLICGLIATVILLIIQLDSGVSDLIKKQ-----ANIH--QEDILLEG---- 83

QY 83 DVEAOATCNHTVMALMSLDRAKAGOGKVVELGEITTLNHHKQDASAEVERLRRENO 142
   ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 84 -----ILAQRSEKS-AQSOKELKEMIEITLAHKIDSKSKIMELHRL 127

QY 143 VLSVRADKKRY--PSSD 159
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 128 NLQEYLKEAANYSGPCPD 146

RESULT        6
US-09-352-302-4
Sequence 4, Application US/09352302
Patent No. 6197937
GENERAL INFORMATION:
APPLICANT: Sawamura, Tatsuya
APPLICANT: Masaki, Tomoo
TITLE OF INVENTION: Modified Low-Density Lipoprotein Receptor
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: McAulay Fisher Nissen Goldberg & Kiel
STREET: 261 Madison Avenue
CITY: New York

```

STATE: NY
COUNTRY: USA
ZIP: 10016-2391
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/352,302
FILING DATE: 12-JUL-1999
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-321705
FILING DATE: 30-NOV-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 7-214206
FILING DATE: 31-JUL-1995
ATTORNEY/AGENT INFORMATION:
NAME: Goldberg, Jules E
REGISTRATION NUMBER: 24408
REFERENCE/DOCKET NUMBER: JG-VY-4363PCT/D
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 986-4090
TELEFAX: 212 818-9479
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 273 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-352-302-4

Query Match 11.2%; Score 99.5; DB 4; Length 273;
Best Local Similarity 25.9%; Pred. No. 0.0069;
Matches 36; Conservative 25; Mismatches 41; Indels 37; Gaps 5;

QY 23 LIGIGILVLLIIVILGPIITIKANSEACRDGLRAVMECRVTHLLOELTEAOKGFQ 82
DB 43 VLGIGLVVLLIILQLSQVSDLIKQQ-----ANTH--QEDIDEG----- 83

QY 83 DVNAQATCNHTYMALMASIDAKAGQKVEELEGITTLNHLQDASAEVRLRENOVL 142
DB 84 -----ILAQRRSEK-AQESQKELKEWETLAKHDKSKIMELHRONL 127

143 VLSVRIADKKY--PSSOD 159
128 NLOEVILKEANYSGPCPD 146

RESULT 7
US-08-533-306A-4
Sequence 4, Application US/08533306A
Patent No. 5837457
GENERAL INFORMATION:
APPLICANT: Liu, Pu
APPLICANT: Collins, Francis S.
APPLICANT: Siciliano, Michael J.
APPLICANT: Claxton, David
TITLE OF INVENTION: Markers for Detection of Chromosome 16
TITLE OF INVENTION: Rearrangements
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Harness, Dickey & Pierce, P.L.C.
STREET: P.O. Box 828
CITY: Bloomfield Hills
STATE: MI
COUNTRY: USA
ZIP: 48303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/533,306A
FILING DATE: September 25, 1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Smith, Deann F.
REGISTRATION NUMBER: 36683
REFERENCE/DOCKET NUMBER: 2115-00869COB
TELECOMMUNICATION INFORMATION:
TELEPHONE: (810) 641-1600
TELEFAX: (810) 641-0270
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 885 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-533-306A-4

Query Match 10.7%; Score 95; DB 2; Length 885;
Best Local Similarity 24.5%; Pred. No. 0.11;
Matches 27; Conservative 18; Mismatches 55; Indels 10; Gaps 1;

QY 53 CRGGLRAVMECRVTHLLOELTEAOKGFQDVNAQATCNHTYMALMASID----- 103
DB 215 CSQGERARALNDKVKHLQNEVESYGMLEAEGAKAKLAKDVASLSQDQDQELLE 274

QY 104 -AEKAGQKVEELEGITTLNHLQDASAEVRLRENOVL 152
DB 275 TROKLVNSTRKLEERNSLQDQDEMEAKONLEHISTLNQLSDSK 324

RESULT 8
US-08-742-923A-4
Sequence 4, Application US/08742923A
Patent No. 5869611
GENERAL INFORMATION:
APPLICANT: Liu, Pu
APPLICANT: Collins, Francis S.
APPLICANT: Siciliano, Michael J.
APPLICANT: Claxton, David
TITLE OF INVENTION: Markers for Detection of Chromosome 16
TITLE OF INVENTION: Rearrangements
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Harness, Dickey & Pierce, P.L.C.
STREET: P.O. Box 828
CITY: Bloomfield Hills
STATE: MI
COUNTRY: USA
ZIP: 48303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/742,923A
FILING DATE: No. 5869611ember 1, 1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Smith, Deann F.
REGISTRATION NUMBER: 36683
REFERENCE/DOCKET NUMBER: 2115-00869DVC
TELECOMMUNICATION INFORMATION:
TELEPHONE: (810) 641-1600
TELEFAX: (810) 641-0270
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 885 amino acids


```

1 APPLICANT: Cocks, Benjamin G.
2 APPLICANT: Goll, Surya K.
3 APPLICANT: Hillman, Jennifer L.
4 TITLE OF INVENTION: NOVEL HUMAN C-TYPE LECTIN
5 NUMBER OF SEQUENCES: 5
6 CORRESPONDENCE ADDRESS:
7 ADDRESSEE: Incyte Pharmaceuticals, Inc.
8 STREET: 3174 Porter Drive
9 CITY: Palo Alto
10 STATE: CA
11 COUNTRY: US
12 ZIP: 94304
13 COMPUTER READABLE FORM:
14 MEDIUM TYPE: Diskette
15 COMPUTER: IBM Compatible
16 OPERATING SYSTEM: DOS
17 SOFTWARE: FASTSEQ Version 1.5
18 CURRENT APPLICATION DATA:
19 APPLICATION NUMBER: US/09/113,788
20 FILING DATE:
21 PRIOR APPLICATION DATA:
22 APPLICATION NUMBER: 08/688,342
23 FILING DATE:
24 ATTORNEY/AGENT INFORMATION:
25 NAME: Billings, Lucy J.
26 REGISTRATION NUMBER: 36,749
27 REFERENCE/DOCKET NUMBER: PF-0095-1 CIP
28 TELECOMMUNICATION INFORMATION:
29 TELEPHONE: 415-855-0555
30 TELEFAX: 415-845-4166
31 INFORMATION FOR SEQ ID NO: 4:
32 SEQUENCE CHARACTERISTICS:
33 LENGTH: 292 amino acids
34 TYPE: amino acid
35 STRANDEDNESS: single
36 TOPOLOGY: linear
37 MOLECULE TYPE: peptide
38 IMMEDIATE SOURCE:
39 LIBRARY: GenBank
40 CLONE: 1235724
41 US-09-113-788-4
42
43 Query Match 10.4%; Score 92.5; DB 2; Length 292;
44 Best Local Similarity 24.6%; Pred. No. 0.043; Indels 31; Gaps
45 Matches 32; Conservative 28; Mismatches 39;
46
47 20 CKL--LGIGILVLLIVLVPLIIFTKANSEACROGLRAVMECRNTYHLOOELTEA 77
48 | || ||:||:||:||:|| ||:||:||:||:|| ||:||:||:||:|| |||||
49 39 CHLLSLTGIGILLVITICVF-----QNSKFORDLVTLRTPDSNFTSNMVAEI--- 87
50
51 Oy 78 OKGFODVERQAATCNHTYVALMASLDAAEKAGQ-----KKVEELGEITTL 123
52 | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
53 Db 88 ----CAITSOGSSLEETITASLKAEVGEGRQAVHSEMILRVQQLVDLKLRTLCOVATL 143
54 | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
55 Oy 124 NKKIODASAE 133
56 | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
57 Db 144 NNNGEASTE 153
58
59 RESULT 12
60 US-09-111-470-4
61 Sequence 4, Application US/09111470
62 Patent No. 6277959
63 GENERAL INFORMATION:
64 APPLICANT: Valladeau, Jenny
65 APPLICANT: Ravel, Odile
66 APPLICANT: Bates, Elizabeth E.M.
67 APPLICANT: Ford, John
68 APPLICANT: Saeland, Sem
69 APPLICANT: Lebecque, Serge J.E.
70 TITLE OF INVENTION: Mammalian Membrane Protein Genes;
71 Related Reagents
72

```

```

CORRESPONDENCE SEQUENCE: 11
ADDRESS: DMAX Research Institute
STREET: 901 California Avenue
CITY: Palo Alto
STATE: California
COUNTRY: USA
ZIP: 94304-1104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/111,470
FILING DATE: 08-JUL-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/053,080
FILING DATE: 09-JUL-1997
ATTORNEY/AGENT INFORMATION:
NAME: Ching, Edwin P.
REGISTRATION NUMBER: 34,090
REFERENCE/DOCKET NUMBER: SP0695
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650)852-9196
TELEFAX: (650)496-1200
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 316 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-111-470-4

Query Match      10.1%; Score 89.5; DB 4; Length 316;
Best Local Similarity 27.8%; Pred. No. 0.1;
Matches 30; Conservative 20; Mismatches 41; Indels 17; Gaps 3.

OY    20 CKL-LGIGILVLLIIVLGVPILITFIKANSEACRGLRAVMECRNVTILLQDELTEA 77
      ||| |||||:::||:|||||
DB    39 CHLLSLGIGLLLVLCVGF-----QNSKFRDLVTLRTFDNFSNTSVTAEL--- 87

OY    78 QKGPDVERQAATCNHTYMALMASIDAERAGQGKKVELEGGITTLNH 125
      |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB    88 ----ALTSGSSLETETIASLKAEEVGFKROERAGVSELDEHTTOKAH 131

RESULT 13
US-09-310-187A-1
Sequence 1, Application US/09310187A
Patent No. 6358751
GENERAL INFORMATION:
APPLICANT: Benichou, Gilles
APPLICANT: Fedoseyeva, Eugenia
TITLE OF INVENTION: Involvement of Autoantigens in Cardiac
FILE REFERENCE: UCSF-090
CURRENT APPLICATION NUMBER: US/09/310,187A
CURRENT FILING DATE: 1999-05-12
NUMBER OF SEQ ID NOS: 3
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 1939
TYPE: PRT
ORGANISM: Homo sapiens
US-09-310-187A-1

Query Match      10.1%; Score 89.5; DB 4; Length 1939;
Best Local Similarity 24.4%; Pred. No. 1.3;
Matches 33; Conservative 28; Mismatches 51; Indels 23; Gaps 5;
```


RESULT 15
US-08-446-201-4
; Sequence 4, Application US/08446201B

Search completed: June 4, 2002, 15:26:57
Job time: 130 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2002, 15:22:47 ; Search time 30.06 Seconds

(without alignments)
665.113 Million cell updates/sec

Title: US-09-828-217-1

Perfect score: 889
Sequence: 1 MASTSYDYCVPMEDGDKRC.....SSAAPQLLVILGLSLTLQ 180

Scoring table:
BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 747574 seqs, 11073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : A.GeneSeq_032802:*
1: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT:*
21: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT:*
22: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	889	100.0	180	AA1980.DAT	Human membrane pol
2	889	100.0	180	AA1981.DAT	Protein recognised
3	889	100.0	180	AA1982.DAT	Human HMI.24 anti
4	889	100.0	180	AA1983.DAT	Soluble HMI.24 ant
5	889	100.0	180	AA1984.DAT	HMI.24 antigenic p
6	889	100.0	180	AA1985.DAT	Potential for an
7	889	100.0	180	AA1986.DAT	Bst-2 protein. Mu
8	889	100.0	180	AA1987.DAT	Human HMI.24 anti
9	889	100.0	180	AA1988.DAT	Bone marrow stroma
10	889	100.0	180	AA1989.DAT	Human HMI.24 prote
11	889	100.0	180	AA1990.DAT	

12	889	100.0	193	AA1991.DAT	Human colon cancer
13	882	99.2	180	AA1992.DAT	Protein bound by A
14	843.5	94.9	197	AA1993.DAT	Protein encoded by A
15	649.5	73.1	143	AA1994.DAT	HMI.24 antigenic p
16	649	73.0	132	AA1995.DAT	Soluble HMI.24 ant
17	649	73.0	161	AA1996.DAT	HMI.24 antigenic p
18	579	65.1	147	AA1997.DAT	HMI.24 antigenic p
19	574.5	64.6	126	AA1998.DAT	HMI.24 antigenic p
20	473.5	53.3	155	AA1999.DAT	Human HMI.24 anti
21	104.5	11.8	756	AA2000.DAT	Human ORF2369
22	102.5	11.5	1484	AA2001.DAT	Polypeptide #2 for
23	102.5	11.5	270	AA2002.DAT	Canine ribosome re
24	99.5	11.2	270	AA2003.DAT	Low density lipopr
25	99.5	11.2	270	AA2004.DAT	Bovine LDL recepto
26	99.5	11.2	273	AA2005.DAT	Bovine LOX-1 polyp
27	99.5	11.2	273	AA2006.DAT	Low density lipopr
28	99	11.1	426	AA2007.DAT	Human ORF1960
29	99	11.1	694	AA2008.DAT	Human ORF1960
30	99	11.1	962	AA2009.DAT	Human transport-as
31	95.5	10.7	909	AA2010.DAT	Human colon cancer
32	95.5	10.7	977	AA2011.DAT	Amino acid sequenc
33	95.5	10.7	977	AA2012.DAT	Human protein SEQ
34	95.5	10.7	1003	AA2013.DAT	Human protein SEQ
35	95.5	10.7	1240	AA2014.DAT	Amino acid sequenc
36	95	10.7	885	AA2015.DAT	AMUL chromosome in
37	94.5	10.6	111	AA2016.DAT	Gene #29 human sec
38	93.5	10.5	140	AA2017.DAT	H. pylori GHPD 363
39	93.5	10.5	1390	AA2018.DAT	Drosophila melanog
40	93.5	10.5	1411	AA2019.DAT	Nucleolar/endosoma
41	93.5	10.5	1886	AA2020.DAT	Rattus norvegicus
42	93	10.5	944	AA2021.DAT	Snail human diagno
43	93	10.5	1972	AA2022.DAT	Smooth muscle myos
44	92.5	10.4	210	AA2023.DAT	Bacterial general
45	92.5	10.4	488	AA2024.DAT	Amino acid sequenc

ALIGNMENTS

RESULT 1
AA1980.DAT
ID AA1980.DAT standard; Protein; 180 AA.
AC AA1980.DAT
DT 06-DEC-1995 (first entry)
DE Human membrane polypeptide for enhancing pre-B cell growth.
XX Rheumatoid arthritis; diagnosis; pre-B cell growth; enhancement.
OS Homo sapiens.
PN WO9510536-A.
PD 20-APR-1995.
PE 14-OCT-1994; 94WO-01P01732.
PR 15-OCT-1993; 93JP-0281622.
PA (HIRA/) HIRANO T.
XX Hirano T, Kaisho T;
XX WPI; 1995-161738/21.
XX N-PSDB; AAQ89606.
PT Recombinant membrane protein enhancing pre-B cell growth - and
PT monoclonal antibody recognising it and useful in the diagnosis of
PT Rheumatoid arthritis
PS Claim 1; Page 29-30; 40pp; Japanese.

XX A membrane polypeptide has been isolated from rheumatoid arthritis
 CC patients. The polypeptide supports pre-B-cell growth and is useful
 CC as an antigen for generating monoclonal antibodies for diagnosing
 CC rheumatoid arthritis.

XX Sequence 180 AA;

Query Match 100.0%; Score 889; DB 16; Length 180;
 Best Local Similarity 100.0%; Pred. No. 6, 2e-78;
 Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASTSYDCYCPVPMEDGDRCKLLGGIGILVLLITVILGVPILITFTKANSEACDGLRAY 60
 Db 1 mastsydcyrcvpmcdgdkrcklilgylvlvllilgvpiliflkanseacrdglrav 60
 QY 61 MECRNTHLLQOELTEAKQKFOVEAQAATCNFTVAMALSLDAEKAGQKVEELGEI 120
 61 mecrnthlllqgelteakqkfvgveaqaatcnhtvamlsldeakagqkveelegei 120
 QY 121 TTNLHKLDASAEYERLRRENOVLSVRIADKKYPPSSQDSSAAAPQLLIVLIGLSALLQ 180
 Db 121 tlnhklqdasaeveylrrrenqvlsvriadkkyppssqdsaaapqlilvlliglsallq 180

RESULT 2
 AAW65771 standard; Protein; 180 AA.

XX AAW65771;
 AC 20-NOV-1998 (first entry)
 DT Protein recognised by anti-human HML.24 antibody.
 DE
 XX Cytotoxic antibody; anti-human HML.24; lymphocytic tumours;
 KM pre-B lymphoma; Burkitt's lymphoma; T-cell tumour; B-cell tumour.
 OS Homo sapiens.
 XX MO9835698-A1.
 PN 20-AUG-1998.
 PD 12-FEB-1998; 98WO-JP00568.
 PF 12-FEB-1997; 97JP-0041410.
 XX (CHUS) CHUGAI SEIYAKU KK.
 PA Koishihara Y, Yoshimura Y;
 PI WPI: 1998-456869/39.
 DR N-PSDB; AAV07579.
 XX Treatment of lymphocytic tumours using cytotoxic antibody - binding
 PT to specific antigen such as HML.24 and effective against T-cell
 PT tumours and B-cell tumours other than myeloma
 XX
 PS Claim 1; Page 44-45; 82pp; Japanese.

CC The protein having the amino acid sequence below is bound specifically
 CC by a cytotoxic antibody which can be used in the treatment of lymphocytic
 CC tumours, including T-cell tumours and B-cell tumours other than myeloma.
 CC The antibody is preferably monoclonal and has ADCC or CDC type
 CC cytotoxicity. It may be chimeric or humanised, and preferably contains
 CC a human antibody constant region C gamma (such as C gamma 1 or C
 CC gamma 3). A preferred antibody is an anti-human HML.24 antibody or an
 CC antibody which binds to an epitope recognising anti-human HML.24
 CC antibody. The cytotoxic antibody is useful in the treatment of
 CC lymphocytic tumours such as acute or chronic B lymphocytic leukaemia,
 CC pre-B lymphoma, Burkitt's lymphoma, or acute or chronic T lymphocytic

CC Leukemia.

XX Sequence 180 AA;

Query Match 100.0%; Score 889; DB 19; Length 180;
 Best Local Similarity 100.0%; Pred. No. 6, 2e-78;
 Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASTSYDCYCPVPMEDGDRCKLLGGIGILVLLITVILGVPILITFTKANSEACDGLRAY 60
 Db 1 mastsydcyrcvpmcdgdkrcklilgylvlvllilgvpiliflkanseacrdglrav 60
 QY 61 MECRNTHLLQOELTEAKQKFOVEAQAATCNFTVAMALSLDAEKAGQKVEELGEI 120
 Db 61 mecrnthlllqgelteakqkfvgveaqaatcnhtvamlsldeakagqkveelegei 120
 QY 121 TTNLHKLDASAEYERLRRENOVLSVRIADKKYPPSSQDSSAAAPQLLIVLIGLSALLQ 180
 Db 121 tlnhklqdasaeveylrrrenqvlsvriadkkyppssqdsaaapqlilvlliglsallq 180

RESULT 3
 AAW62207 standard; Protein; 180 AA.

XX AAW62207;
 AC 21-SEP-1998 (first entry)
 DT Humanised anti-HML.24 antibody polypeptide.
 DE
 XX Mouse: human; humanised; anti-HML.24 antibody; myeloma; FR; CDR;
 KM framework region; complementarity determining region; antigenicity.
 XX
 OS Synthetic.
 OS Mus sp.
 XX Homo sapiens.
 XX MO9814580-A1.
 PN 09-APR-1998.
 PD 03-OCT-1997; 97WO-JP03553.
 PF 04-OCT-1996; 96JP-0264756.
 PR (CHUS) CHUGAI SEIYAKU KK.
 PA Koishihara Y, Kosaka M, Ohtomo T, Ono K, Tsuchiya M;
 PI Yoshimura Y;
 PI WPI: 1998-286421/25.
 DR N-PSDB; AAV39359.
 XX Humanised anti-HML.24 antibody - for treatment of myeloma
 PT
 PS Claim 81; Page 151-153; 210pp; Japanese.

CC A humanised anti-HML.24 antibody has been developed which comprises
 CC human L and H chain C regions, and L and/or H chain V regions
 CC containing material originating in mouse anti-HML.24 antibody. The V
 CC regions contain framework (FR) regions of human origin and
 CC complementarity determining regions (CDR) of mouse origin, leading to
 CC a reshaped humanised antibody. The C regions are human Ck (L-chain) and
 CC human C gamma (especially C gamma 1) (H-chain). The FR regions of the
 CC L chain V region are derived from human subtype HSG1 (e.g. from human
 CC antibody R61) and the FR regions of the H chain V region are derived
 CC from human subtype HSG1 (e.g. FR1-3 from human antibody H63 and FR4
 CC from human antibody JH6). The present sequence represents an antibody
 CC polypeptide from the present invention. The antibodies are used for
 CC the treatment of myeloma, especially by injection, intravenously,
 CC intramuscularly or subcutaneously. The antibodies are used at 0.01-1000

CC (especially 5-100) mg/kg body weight. The humanised antibody has low
 CC antigenicity and is therefore effective therapeutically in humans.
 XX
 SQ Sequence 180 AA:

Query Match 100.0%; Score 889; DB 19; Length 180;
 Best Local Similarity 100.0%; Pred. No. 6.2e-78;
 Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASTSYDYGKRVPMEDGDKRCKLLGIGILVLLIIVILGVPILITFTKANSEACRDLRAY 60
 DB 1 mastsydygcrvpmedgdkrccklllgigilvllilvllgvpiliftlkansaeacrdglrav 60
 OY 61 MECRANTHLLQOELTFPAOKGFODVEAOAATCNHTVMAIMASIDAEKAGQKVEELGET 120
 DB 61 mecrnvtlllqgelteaqg9fgdveaqaatcnltvmaimasidaekagqkveeleget 120
 OY 121 TTLNHLKODASAEVERLRRENOVL SVRIADKKRYPPSSODSSSAAPOLLIVLGLSALLQ 180
 DB 121 tllnhklqdasaeeverllrrengvlsvriadkkryppsgdsssaapqllilvllglsallq 180

RESULT 4

AA33202
 ID AAY33202 standard; Protein; 180 AA.

AC AAY33202;
 DT 22-NOV-1999 (first entry)
 XX
 DE Human HML.24 antigenic protein.

KW Antigenic protein; HML.24; splice variant; promoter; antineumatic;
 KW antiarthritic; bone marrow; tumour cell; drug development; treatment;
 KW myeloma; rheumatoid arthritis; human.

OS Homo sapiens.

PN WO9943803-A1.

PD 02-SEP-1999.

PF 25-FEB-1999; 99WO-JP00884.

PR 25-FEB-1998; 98JP-0060617.

PR 24-MAR-1998; 98JP-0093883.

PA (CHUS) CHUGAI SEIYAKU KK.

PI Ohtomo T, Tsuchiya M, Koishihara Y, Kosaka M;

DR WPI; 1999-550869/46.

DR N-PSDB; AA209726.

XX Genomic DNA encoding HML.24 antigen protein as well as splicing
 PT variants, useful e.g. in development of drugs for treating myeloma and
 PT rheumatoid arthritis

PS Example 1; Fig 1-2; 83pp; Japanese.

CC This invention describes a novel human antigenic protein, HML.24,
 CC its encoding nucleic acid, splice variants and promoter region. The
 CC products of the invention have antirheumatic and antiarthritic activity.
 CC The DNA of the invention is isolated from bone marrow tumour cells,
 CC which can be used to study the expression of HML.24 antigen, promoter
 CC activity of its promoter region, and in development of drugs in treating
 CC e.g. myeloma and rheumatoid arthritis. This sequence represents the
 CC human HML.24 antigenic protein described in the invention.

SQ Sequence 180 AA:

Query Match 100.0%; Score 889; DB 20; Length 180;
 Best Local Similarity 100.0%; Pred. No. 6.2e-78;
 Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASTSYDYGKRVPMEDGDKRCKLLGIGILVLLIIVILGVPILITFTKANSEACRDLRAY 60
 DB 1 mastsydygcrvpmedgdkrccklllgigilvllilvllgvpiliftlkansaeacrdglrav 60
 OY 61 MECRANTHLLQOELTFPAOKGFODVEAOAATCNHTVMAIMASIDAEKAGQKVEELGET 120
 DB 61 mecrnvtlllqgelteaqg9fgdveaqaatcnltvmaimasidaekagqkveeleget 120
 OY 121 TTLNHLKODASAEVERLRRENOVL SVRIADKKRYPPSSODSSSAAPOLLIVLGLSALLQ 180
 DB 121 tllnhklqdasaeeverllrrengvlsvriadkkryppsgdsssaapqllilvllglsallq 180

RESULT 5

AA32765
 ID AAY32765 standard; Protein; 180 AA.

AC AAY32765;
 DT 25-OCT-1999 (first entry)
 XX
 DE Soluble HML.24 antigenic protein sequence.

KW HML.24; antigenic protein; antibody; immunoassay; diagnosis; cancer;
 KW immune disorder; immunotherapy monitoring.
 XX
 OS Homo sapiens.

PN WO9943703-A1.

PD 02-SEP-1999.

PF 25-FEB-1999; 99WO-JP00885.

PR 25-FEB-1998; 98JP-0060613.

PA (CHUS) CHUGAI SEIYAKU KK.

PI Koishihara Y, Ozaki Y;

DR WPI; 1999-518836/43.

DR N-PSDB; AA210917.

PT Immunoassay of anti-HML.24 antibody or soluble HML.24 antigen,
 PT useful for diagnosis of immune disorders and cancer

PS Example 3; Fig 14-15; 138pp; Japanese.

CC This sequence represents a human soluble HML.24 antigenic protein. The
 CC invention relates to an immunochemical assay of anti-HML.24 antibody by
 CC use of a soluble HML.24 antigenic protein, or an immunochemical assay of
 CC the soluble antigen by use of the antibody. The immunoassay of the HML.24
 CC antigen or antibody is useful for diagnosis of immune disorders and
 CC cancer, for monitoring of anti-HML.24 antibody immunotherapy, and for
 CC assay of the antibody or antigen for investigative purposes, in
 CC biological samples such as blood, serum, urine, milk, synovial fluid or
 CC microorganism culture media. The method is sensitive down to 500 pg/ml
 CC antibody.

SQ Sequence 180 AA:

Query Match 100.0%; Score 889; DB 20; Length 180;
 Best Local Similarity 100.0%; Pred. No. 6.2e-78;
 Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASTSYDYGKRVPMEDGDKRCKLLGIGILVLLIIVILGVPILITFTKANSEACRDLRAY 60
 DB 1 mastsydygcrvpmedgdkrccklllgigilvllilvllgvpiliftlkansaeacrdglrav 60

```
QY 61 MEGRNTHLLQOELTEAOKGFQDVEAQAATCNHTVMAIAMSLEAKAGOKKVEELEGEI 120
DB 61 mecrnvtlllqgelteagkfqdvaaqaatcnhtvmaiasldeakagqkveelegei 120
QY 121 TLTNHRKIQDASAEVERLRRENOVLSVRIDAKKYPSSODSSAAAPQLLIVLGLSALLQ 180
DB 121 tltnhkIQdasaeverlrrenqvlsvrIdackkypssqdsassaapqllivlglSallq 180

RESULT 6
AAV02576
ID AAV02576 standard; Protein; 180 AA.
XX
AC AAV02576;
XX
DT 16-JUL-1999 (first entry)
XX
XX HMI.24 antigenic protein.
XX
XX Reconstituted human antibody; peptide antigen HMI.24; framework region;
XX complementary determining region; CDR; anti-HMI.24 antibody; myeloma;
XX humanised antibody.
XX
XX Homo sapiens.
XX
XX WO918212-A1.
XX
XX 15-APR-1999.
XX
XX 02-OCT-1998; 98WO-JP04469.
XX
XX 03-OCT-1997; 97JP-0271726.
XX
XX (CHUS ) CHUGAI SEIYAKU KK.
XX
XX Tsuchiya M;
XX
XX WPI; 1999-277273/23.
XX
XX N-PSDB; AAX59485.
XX
XX Reconstituted human antibody useful in the treatment of myeloma
XX
XX Disclosure; Page 131-133; 256pp; Japanese.
XX
XX The specification describes a reconstituted human antibody recognizing
XX the peptide antigen HMI.24. This human antibody contains natural human
XX framework regions modified by amino acid substitutions to provide
XX homogeneity with a previously designed framework region (which may
XX arise from a human or non-human source); and complementary determining
XX regions (CDR) derived from a non-human anti-HMI.24 antibody. The
XX reconstituted antibody is useful in the treatment of diseases in which
XX the surface antigen HMI.24 is implicated such as myeloma. The present
XX sequence represents HMI.24 antigenic protein.
XX
XX Sequence 180 AA;

Query Match 100.0%; Score 889; DB 20; Length 180;
Best Local Similarity 100.0%; Pred. No. 6.2e-78;
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASTSYDYCRVPMEDGDKRCKLLIGIILVLLIIVLGVPLIIFTTKANSEACRDLIRAV 60
DB 1 mastsydyrcvpmegdckrcklllgilvlllIvllgvplllfttkanseacrDglrav 60
QY 61 MEGRNTHLLQOELTEAOKGFQDVEAQAATCNHTVMAIAMSLEAKAGOKKVEELEGEI 120
DB 61 mecrnvtlllqgelteagkfqdvaaqaatcnhtvmaiasldeakagqkveelegei 120
QY 121 TLTNHRKIQDASAEVERLRRENOVLSVRIDAKKYPSSODSSAAAPQLLIVLGLSALLQ 180
DB 121 tltnhkIQdasaeverlrrenqvlsvrIdackkypssqdsassaapqllivlglSallq 180
```

```
RESULT 7
AAV05484
ID AAV05484 standard; Protein; 180 AA.
XX
AC AAV05484;
XX
DT 07-JUL-1999 (first entry)
XX
XX Potentiator for antibody against lymphoid tumour.
XX
XX DE
XX
XX Antibody potentiator; lymphoid tumour; lymphoma; cytotoxic antibody;
XX multiple myeloma; acute B-lymphocytoma; chronic B-lymphocytoma;
XX pre-B lymphoma; Burkitt's lymphoma; acute T-lymphocytoma; therapy;
XX chronic T-lymphocytoma; PNTL.
XX
XX Homo sapiens.
XX
XX WO918997-A1.
XX
XX 22-APR-1999.
XX
XX 14-OCT-1998; 98WO-JP04645.
XX
XX 05-AUG-1998; 98JP-0222024.
XX
XX 14-OCT-1997; 97JP-0280759.
XX
XX (CHUS ) CHUGAI SEIYAKU KK.
XX
XX Koishihara Y, Kosaka M;
XX
XX WPI; 1999-277447/23.
XX
XX N-PSDB; AAX36561.
XX
XX Claim 1; Page 37-38; 62pp; Japanese.
XX
XX This sequence represents a potentiator for an antibody against
XX lymphoid tumour.
XX
XX The invention relates to a method for the treatment of lymphoma, in which
XX a cytotoxic antibody is potentiated by administration of a biological
XX response modifier. The method can be used for treatment of lymphomas and
XX multiple myelomas which are resistant to conventional treatment, such as
XX acute B-lymphocytoma, chronic B-lymphocytoma, pre-B lymphoma, Burkitt's
XX lymphoma, acute T-lymphocytoma, chronic T-lymphocytoma, and PNTL.
XX
XX Sequence 180 AA;

Query Match 100.0%; Score 889; DB 20; Length 180;
Best Local Similarity 100.0%; Pred. No. 6.2e-78;
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASTSYDYCRVPMEDGDKRCKLLIGIILVLLIIVLGVPLIIFTTKANSEACRDLIRAV 60
DB 1 mastsydyrcvpmegdckrcklllgilvlllIvllgvplllfttkanseacrDglrav 60
QY 61 MEGRNTHLLQOELTEAOKGFQDVEAQAATCNHTVMAIAMSLEAKAGOKKVEELEGEI 120
DB 61 mecrnvtlllqgelteagkfqdvaaqaatcnhtvmaiasldeakagqkveelegei 120
QY 121 TLTNHRKIQDASAEVERLRRENOVLSVRIDAKKYPSSODSSAAAPQLLIVLGLSALLQ 180
DB 121 tltnhkIQdasaeverlrrenqvlsvrIdackkypssqdsassaapqllivlglSallq 180

RESULT 8
AAV07250
ID AAV07250 standard; Protein; 180 AA.
XX
```

AC AAY07250;
XX
DT 06-JUL-1999 (first entry)
XX
DE BST-2 protein.
XX
KW Mouse; BST-2; monoclonal antibody; RS38; myeloma; cytotoxic activity.
XX
OS Mus sp.
XX JP11092399-A.
XX
PD 06-APR-1999.
XX
PE 24-SEP-1997; 97JP-0274960.
XX
PR 24-SEP-1997; 97JP-0274960.
XX
PA (CHUS) CHUGAI PHARM CO LTD.
XX
NR WPI: 1999-283503/24.
XX N-PSDB: AAX29996.
XX
PT An agent for treating myeloma -includes an antibody and has
XX cytotoxic activity
XX
PS Claim 1: Page 10; 13pp; Japanese.
XX
CC This sequence represents the mouse BST-2 protein which is used to raise
CC antibodies, especially the monoclonal antibody RS38. The antibody can be
CC used in compositions to treat myelomas when the antibody is associated
CC with a cytotoxic activity.
XX
SO Sequence 180 AA:

Query Match 100.0%; Score 889; DB 20; Length 180;
Best Local Similarity 100.0%; Pred. No. 6.2e-78;
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASTSYDYCRVPMEDGDKRCKLLGIGILVLLIIVLGVPLIIFITKANSEACRDLRAV 60
DB 1 mastsydycrvpmmedgdkrcklllgilvllivllgvplllftlkansacrdglrav 60
QY 61 MECRWVTHLLOELTFAQKGFQDVEAQAATCNHTVWALMASIDAERAGQKVEELEGEL 120
DB 61 mecrwvthllqelteaagqgfdveaqaatcnhtvwalmasidaeragqkveelegel 120
QY 121 TTLNHLKLDASAEVERLRENOVLSVRIADKKRYPPSSQSSAAPQLLIVLLGSLALQ 180
DB 121 ttlnhklldasaeverlrrenovlsvriadkkryppssqssaaapqllivllglsallq 180

RESULT 9
ID AAY53273
AC AAY53273 standard; Protein; 180 AA.
XX
AC AAY53273;
XX
DT 21-JUL-2000 (first entry)
XX
DE Human HML.24 antigen protein sequence SEQ ID NO:2.
XX
KW Human; HML.24 antigen protein; detection; plasmocytoma;
KW multiple myeloma; plasmocytic leukemia; extramedullary plasmocytoma;
KW multiple plasmocytoma; asymptomatic myeloma.
XX
OS Homo sapiens.
XX
PN WO200017395-A1.
XX
PD 30-MAR-2000.
XX

PF 20-AUG-1999; 99WO-JP04502.
XX
PR 18-SEP-1998; 98JP-0264593.
XX
PA (CHUS) CHUGAI SEIYAKU KK.
XX
PI Kawai S, Koshidara Y, Kosaka M;
XX
DR WPI: 2000-283616/24.
DR N-PSDB: AAA13654.
XX
PT Detection or measurement of plasmocytomas, applicable for early
PT diagnosis of e.g. multiple myeloma and plasmocytic leukemia, using a
PT polynucleotide which is expressed specifically or strongly in
PT plasmocytomas
XX
PS Disclosure; Page 15-16; 20pp; Japanese.
XX
CC A method has been developed for detecting or measuring plasmocytomas in
CC a sample at an early stage of disease development. The method comprises
CC amplifying a polynucleotide which is expressed specifically or strongly
CC in plasmocytomas before quantifying the amplification product by
CC comparing with results obtained with a control sample. The method is for
CC detecting or measuring plasmocytomas, applicable for early diagnosis of
CC e.g. multiple myeloma, plasmocytic leukemia, isolated plasmocytoma,
CC extramedullary plasmocytoma, multiple plasmocytoma obtained from smoking
CC or asymptomatic myeloma. The present sequence represents human HML.24
CC antigen protein, which is expressed in plasmocytomas and so can be used
CC in the method of the invention.
XX
SO Sequence 180 AA:

Query Match 100.0%; Score 889; DB 21; Length 180;
Best Local Similarity 100.0%; Pred. No. 6.2e-78;
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASTSYDYCRVPMEDGDKRCKLLGIGILVLLIIVLGVPLIIFITKANSEACRDLRAV 60
DB 1 mastsydycrvpmmedgdkrcklllgilvllivllgvplllftlkansacrdglrav 60
QY 61 MECRWVTHLLOELTFAQKGFQDVEAQAATCNHTVWALMASIDAERAGQKVEELEGEL 120
DB 61 mecrwvthllqelteaagqgfdveaqaatcnhtvwalmasidaeragqkveelegel 120
QY 121 TTLNHLKLDASAEVERLRENOVLSVRIADKKRYPPSSQSSAAPQLLIVLLGSLALQ 180
DB 121 ttlnhklldasaeverlrrenovlsvriadkkryppssqssaaapqllivllglsallq 180

RESULT 10
ID ABB50295
AC ABB50295 standard; Protein; 180 AA.
XX
AC ABB50295;
XX
DT 08-FEB-2002 (first entry)
XX
DE Bone marrow stromal antigen (BST-2) ovarian tumour marker protein, #80.
XX
KW Ovarian tumour marker gene; human; overexpression; upregulation;
KW epithelial tumour; cancer; diagnosis; prognosis; disease monitoring;
KW serous cystadenocarcinoma; mucinous cystadenocarcinoma;
KW mucinous cystadenoma; borderline mucinous tumour; endometrioid carcinoma;
KW undifferentiated carcinoma; clear cell adenocarcinoma; cystadenofibroma;
KW adenofibroma; Brenner tumour; serial analysis of gene expression; SAGE;
KW immune response pathway; cell proliferation regulation; protein folding;
KW membrane localised; secreted; therapeutic target; cytostatic;
KW gene therapy; vaccine.
XX
OS Homo sapiens.
XX

AC AAG73947;
XX
DT 03-SEP-2001 (first entry)
XX
DE Human colon cancer antigen protein SEQ ID NO:4711.
XX
KW Human; colon cancer; colon cancer antigen; diagnosis; detection;
KM colorectal carcinoma; chromosome 19.
XX
OS Homo sapiens.
XX
PN WO200122920-A2.
XX
PD 05-APR-2001.
XX
PE 28-SEP-2000; 2000WO-US26524.
XX
PR 29-SEP-1999; 99US-0157137.
PR 03-NOV-1999; 99US-0163280.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Ruben SM, Barash SC, Birse CE, Rosen CA;
XX
DR WPI; 2001-235357/24.
DR N-PSDB; AAH33378.
XX
PT Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
PT useful for preventing, diagnosing and/or treating colorectal cancers -
XX
PS Claim 11; Page 6512-6513; 9803PP; English.
XX
XX AAH32943 to AAH37195 and AAG73514 to AAG7788 represent human colon
CC cancer-associated nucleic acid molecules (N) and proteins (P), where
CC the proteins are collectively known as colon cancer antigens. The colon
CC cancer antigens have cytostatic activity and can be used in gene
CC therapy and vaccine production. N and P may be used in the prevention,
CC diagnosis and treatment of diseases associated with inappropriate P
CC expression. For example, N and P may be used to treat disorders
CC associated with decreased expression by rectifying mutations or deletions
CC in a patient's genome that affect the activity of P by expressing
CC inactive proteins or to supplement the patient's own production of P.
CC Additionally, N may be used to produce the colon cancer-associated Ps,
CC by inserting the nucleic acids into a host cell and culturing the cell
CC to express the proteins. N and P can be used in the prevention, diagnosis
CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
CC and AAG7789 represent sequences used in the exemplification of the
CC present invention.
CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were
CC missing at time of publication, meaning no sequences are present for
CC SEQ ID NO:1027 to 1052, 7921 and 7922.
XX
SQ Sequence 193 AA;
XX
Query Match 100.0%; Score 889; DB 22; Length 193;
Best Local Similarity 100.0%; Pred. No. 6,8e-78;
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 MASTSYDYCRVPMEDGDKRCKLLGIGILVLLIIVLIGVPLITFTKANSEACRDLRAV 60
DB 14 mastsydycrvpmegdckrcklllgigilvllllypllftkkanseacrdglrav 73
XX
QY 61 MECRNVTHTLLOQELTEAOKGFQDVDEAQAATCNHTVMAALASLDAEKAQOKKVEELEG 120
DB 74 mecrnvthllqelteaogkfqdvdeagaatcnhtvmaasldaekagqkkveelegel 133
XX
QY 121 TLTNHLKLDASAEVRLRENOVLSVRIDAKKRYPPSSODSSAAAPOLLIVLLIGLSALQ 180
DB 134 tltlnhklqdasaeveerlrenovlsvrldakkyppssodssaaapqllivlllglsallq 193
XX
RESULT 13

AAW77292
ID AAW77292 standard; Protein; 180 AA.
XX
XX AAW77292;
AC
XX
DT 14-DEC-1998 (first entry)
XX
DE Protein bound by Anti-HM1.24 antibody.
XX
DE
XX
KW Anti-HM1.24; antibody; lymphocyte activation inhibitor;
KM autoimmune disease; organ transplant; allergy.
XX
OS Homo sapiens.
XX
PN WO9837913-A1.
XX
PD 03-SEP-1998.
XX
PE 27-FEB-1998; 98WO-JP00831.
XX
PR 28-FEB-1997; 97JP-0045663.
XX
PA (CHUS) CHUGAI SEIYAKU KK.
XX
PI Kolshihara Y;
XX
DR WPI; 1998-480937/41.
DR N-PSDB; AAW59114.
XX
PT Lymphocyte activation inhibitor comprises antibodies, particularly
PT anti-HM1.24 antibody - for preventing and treating autoimmune
XX diseases, rejection reactions in organ transplant or allergy
XX
PS Disclosure; Page 38-39; 53pp; Japanese.
XX
CC The Anti-HM1.24 antibody can be used in the production of lymphocyte
CC activation inhibitors. These inhibitors can be used for the prevention
CC and treatment of autoimmune diseases, rejection reactions in organ
CC transplant or allergy. Administration is non-oral, e.g. by intra
CC venous and intramuscular injection, local or systemic.
XX
SQ Sequence 180 AA;
XX
Query Match 99.2%; Score 882; DB 19; Length 180;
Best Local Similarity 99.4%; Pred. No. 3e-77; 1; Indels 0; Gaps 0;
Matches 179; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
QY 1 MASTSYDYCRVPMEDGDKRCKLLGIGILVLLIIVLIGVPLITFTKANSEACRDLRAV 60
DB 1 mastsydycrvpmegdckrcklllgigilvllllypllftkkanseacrdglrav 60
XX
QY 61 MECRNVTHTLLOQELTEAOKGFQDVDEAQAATCNHTVMAALASLDAEKAQOKKVEELEG 120
DB 61 mecrnvthllqelteaogkfqdvdeagaatcnhtvmaasldaekagqkkveelegel 120
XX
QY 121 TLTNHLKLDASAEVRLRENOVLSVRIDAKKRYPPSSODSSAAAPOLLIVLLIGLSALQ 180
DB 121 tltlnhklqdasaeveerlrenovlsvrldakkyppssodssaaapqllivlllglsallq 180
XX
RESULT 14
AAW36951
ID AAW36951 standard; Protein; 197 AA.
XX
XX AAW36951;
AC
XX
DT 12-MAY-1998 (first entry)
XX
DE Protein encoded by clone 0238_1.
XX
KW Human; secreted protein; molecular weight marker; genetic fingerprinting;
KW antibody production; nutritional supplement; therapy; clone 0238_1;
XX

THIS PAGE BLANK (USPTO)